

Quantum Mechanics of a Point Particle in 2+1 Dimensional Gravity

Hans-Jürgen Matschull¹ and Max Welling²
 Instituut voor Theoretische Fysica, Universiteit Utrecht
 Princetonplein 5, 3584 CC Utrecht, The Netherlands

Abstract

We study the phase space structure and the quantization of a pointlike particle in 2+1 dimensional gravity. By adding boundary terms to the first order Einstein Hilbert action, and removing all redundant gauge degrees of freedom, we arrive at a reduced action for a gravitating particle in 2+1 dimensions, which is invariant under Lorentz transformations and a group of generalized translations. The momentum space of the particle turns out to be the group manifold $SL(2)$. Its position coordinates have non-vanishing Poisson brackets, resulting in a non-commutative quantum spacetime. We use the representation theory of $SL(2)$ to investigate its structure. We find a discretization of time, and some semi-discrete structure of space. An uncertainty relation forbids a fully localized particle. The quantum dynamics is described by a discretized Klein Gordon equation.

Outline

The most important property of general relativity in 2+1 dimensions is the lack of local degrees of freedom [1, 2]. There are not enough dimensions to support gravitational waves, which transmit the gravitational interaction. The theory is not completely “void” however. One can introduce so called topological degrees of freedom by choosing a spacetime manifold that is not simply connected. The study of these cosmological models was greatly facilitated by the observation that 2+1-dimensional gravity is equivalent to a Chern-Simons theory of the Poincaré group [3, 4]. The physical degrees of freedom can be described by Wilson loops. Also a great deal of progress was made in the quantization of these models, although it is still not clear whether all approaches result in equivalent quantum theories [5].

Another group of degrees of freedom are usually introduced as punctures in the surface. To these punctures one can attach a parameter, the mass, by adding a particle Lagrangian to the Einstein Hilbert action. It consists of a kinetic term and a constraint, which puts the particle on a mass shell. Because of the topological nature of the interaction, the curvature vanishes everywhere except at the location of the particle. The

¹E-mail: H.J.Matschull@fys.ruu.nl

²E-mail: welling@vision.caltech.edu

effect of the mass is to cut out a wedge from spacetime in such a way that the particle is located at the tip of a cone. For a static particle for instance, the deficit angle of this cone is $\alpha = 8\pi G m$, where m is the mass of the particle. Newton's constant G has the dimension of an inverse mass in three dimensions. Multi particle models are well understood at the level of classical equations of motion [7]. At the quantum level, it is possible to calculate the exact scattering amplitude for two gravitating particles [8], but a consistent multi particle Hilbert space has not been found.

The fact that point particles in 2+1 dimensional gravity are non-trivial objects was demonstrated only recently by 't Hooft [9]. He suggested that the momentum of the particle, which is canonically conjugate to the distance from an arbitrary origin, is actually an angle. This fact was exploited to perform an unusual quantization. He defined the momentum on a sphere and the energy on a circle, resulting in non-commuting operators for the coordinates. The Schrödinger equation became a Dirac equation on some kind of spacetime lattice. This was not the first time that someone proposed to use a curved momentum space however. Back in 1946 Snyder [12] proposed to use de-Sitter or anti-de-Sitter space as momentum space. In any case, the deviation from flat momentum space will not be noticeable at small energy scales. However, at large energy scales, the deviation influences the short distance behaviour in spacetime.

In this article, we want to derive the phase space of a gravitating particle from first principles. We start from the first order Einstein Hilbert action for 2+1 dimensional gravity and add a particle term to it, which is chosen such that the neighbourhood of the particle looks like a cone with a deficit angle of $8\pi G m$. We also have to add a boundary term to the action, in order to render the Lagrangian finite and to obtain the correct equations of motion at spatial infinity. This procedure will, almost automatically, introduce an additional degree of freedom into the system, which can be associated with an observer sitting at infinity. As a result, the gravitational field, which itself has no local degrees of freedom, describes the interaction of the particle with an external observer.

After dividing out the gauge degrees of freedom, we arrive at a finite dimensional reduced phase space, which is spanned by position and momentum coordinates. The position is a vector in Minkowski space, which can be interpreted as the location of the particle measured in the rest frame of the external observer. So far, the description is the same as that of a relativistic point particle in flat Minkowski space. However, the momentum of the particle is not a vector, but an element of the spinor representation of the Lorentz group $SL(2)$. So, the momentum space is indeed curved, but it is not a sphere. Instead, it is locally isomorphic to anti-de-Sitter space. Its cotangent space is the algebra $\mathfrak{sl}(2)$, which is isomorphic to Minkowski space, and thus to the position space of the particle. Hence, all together the phase space becomes the cotangent bundle $T^*SL(2)$.

This has some consequences for the Poisson algebra and the symmetries of the phase space. For example, the components of the position vector will not commute with each other. At the quantum level, this leads to a new kind of uncertainty relation, which states that it is not possible to localize the particle within a region that is smaller than the Planck size. We can also say that the quantized particle lives in a “non-commutative” spacetime. It is important to note that this comes out as a result. We do not start from a non-commutative structure of spacetime or anything like that *a priori*. To see what

such a non-commutative spacetime looks like, we set up a quantum representation that is very similar to the Euclidean angular momentum Hilbert space in three dimensions.

The basic symmetry group that acts on the reduced phase space, and then also on the quantum Hilbert space, turns out to be the Poincaré group. This is not unexpected, because the position coordinate of the particle lives in Minkowski space. However, and this is a somewhat surprising effect, this group is not represented in the ordinary way. In particular, a translation is not represented as a simple shift of the position coordinates, but as a more complicated deformed translation, which also involves the momenta. Only for small momenta, it reduces to an ordinary translation. That the Poincaré group naturally acts on the reduced phase space can also be seen from the fact that the latter is, as a manifold, a special representation thereof, $T^*\text{SL}(2) \simeq \text{SL}(2) \times \mathfrak{sl}(2) \simeq \text{ISL}(2)$, and the natural action is the adjoint action of this group onto itself.

The absence of the ordinary translation symmetry has some consequences for the quantum theory as well. It turns out that there is a distinguished point in spacetime. The light cone emerging from this point divides the quantum spacetime into regions with quite different structures. Space and time are discrete in the causal past and future of this point, whereas in the region that is spacelike separated from this point, space is continuous but time is still discrete. Moreover, there is a smallest distance to this origin, in the sense that the particle cannot be localized in a region of spacetime that is closer than one Planck length from the origin.

Another peculiar effect is that spacetime has an inherent orientation. At the classical level, this shows up in the Poisson brackets of the position coordinates, which involves a Levi Civita symbol. At the quantum level, it manifests itself in the fact that, if we try to localize the particle in a small region, then we have to give it a certain angular momentum as well, which has to point into a fixed direction. Both the distinguished origin of spacetime and its orientation can be considered as parameters of the theory that have to be determined “by experiments”. They can take any value, but it is not possible to remove them completely.

The dynamics of the particle can be described by a mass shell constraint, which is very similar to that of a relativistic point particle. At the quantum level, it becomes a kind of Klein Gordon equation, which is however not a second order differential equation, but a second order difference equation. This is due to the discretization of time, which implies that the time evolution equation has to be a step equation. In the limit where the Planck length goes to zero, it becomes the usual Klein Gordon equation. This is actually true for everything derived from the reduced phase space structure. If we insert the correct physical units, and then switch off the gravitation by making the Planck length (or Newtons constant) small, we recover the usual description of a relativistic point particle.

All these unfamiliar effects can therefore be thought of as gravitational corrections to the relativistic point particle and its quantization, and they result from the combination of quantum mechanics and general relativity. We are aware that this is just a one particle model and the hope is that in a more sophisticated, multi particle model, some of the details will become somewhat clearer. For example, the origin of spacetime might be removed and the absolute position vector of the particle, referring to this origin, be replaced by the relative position between two particles. The shortest distance would then act as a natural cutoff of the theory. Some progress has been made towards

this goal at the level of the classical phase space [13], and a simple model has been worked out in the quantum case [15], but a full understanding is still lacking.

1 The spinor representation of 2+1 gravity

For the purpose of this article, it is most convenient to use the dreibein representation of Einstein gravity. The basic field variables are then the dreibein e_μ^a and the $\mathfrak{so}(1, 2)$ spin connection $\omega_{\mu ab} = -\omega_{\mu ba}$. Here, greek indices μ, ν, \dots are tangent indices of the spacetime manifold \mathcal{M} , and latin indices a, b, \dots , taking the values 0, 1, 2, are three dimensional Minkowski space indices, transforming under the vector representation of the Lorentz group $SO(1, 2)$. They are raised and lowered by the metric $\eta_{ab} = \text{diag}(-1, 1, 1)$, and the Levi-Civita symbol ε^{abc} is defined such that $\varepsilon^{012} = 1$.

A brief introduction into notations and conventions, the structure of the Lorentz group, its spinor representation $SL(2)$, and the associated Lie algebras is given in Appendix A. A special feature of the three dimensional Lorentz algebra is that it is, as a vector space, isomorphic to Minkowski space itself, and thus the vector representation of the Lorentz algebra is equivalent to its adjoint representation. We can use this to convert the dreibein as well as the spin connection into $\mathfrak{sl}(2)$ valued one-forms, defined by

$$e_\mu = e_\mu^a \gamma_a, \quad \omega_\mu = \frac{1}{4} \omega_{\mu ab} \gamma^a \gamma^b. \quad (1.1)$$

The gamma matrices are those introduced in (A.3), and bold letters are used to denote 2×2 matrices. As a general convention, we shall always use the same symbols, either with vector indices attached or in bold face, to denote the same physical quantity in different representations.

Local frames

The dreibein can be considered as a map from the tangent bundle of the spacetime manifold \mathcal{M} into a Minkowski space, or $\mathfrak{sl}(2)$ bundle over \mathcal{M} . A tangent vector $v^\mu(x)$ at some point $x \in \mathcal{M}$ is mapped onto a vector $\boldsymbol{v}(x) = v^\mu(x) e_\mu(x)$ in a “local Lorentz frame”, or simply a “local frame” at x . The scalar product of two such vectors is given by

$$\frac{1}{2} \text{Tr}(\boldsymbol{v} \boldsymbol{w}) = \frac{1}{2} \text{Tr}(e_\mu e_\nu) v^\mu w^\nu = g_{\mu\nu} v^\mu w^\nu. \quad (1.2)$$

This defines the spacetime metric $g_{\mu\nu}$, which is required to be invertible. Equivalently, the dreibein determinant has to be non-zero, and without loss of generality we can restrict it to be positive,

$$e = \frac{1}{12} \varepsilon^{\mu\nu\rho} \text{Tr}(e_\mu e_\nu e_\rho) > 0. \quad (1.3)$$

This condition implies that the dreibein provides an isomorphism between the tangent bundle of \mathcal{M} and the local frame bundle. So, in order to get an invertible metric, the latter has to be isomorphic to the tangent bundle of the given manifold \mathcal{M} . We shall see that, for our special model, the tangent bundle of the spacetime manifold is trivial. Let

us therefore in the following assume that this is the case. The bundle of local frames is then also trivial, and we can consider the dreibein as an $\mathfrak{sl}(2)$ valued one-form on \mathcal{M} . The same holds for the spin connection, which defines the covariant derivative of vectors in the local frames,

$$D_\mu \mathbf{v} = \partial_\mu \mathbf{v} + [\boldsymbol{\omega}_\mu, \mathbf{v}]. \quad (1.4)$$

From this, we derive the following useful quantities. The field strength is defined by

$$[D_\mu, D_\nu] \mathbf{v} = [\mathbf{F}_{\mu\nu}, \mathbf{v}] \quad \Rightarrow \quad \mathbf{F}_{\mu\nu} = \partial_\mu \boldsymbol{\omega}_\nu - \partial_\nu \boldsymbol{\omega}_\mu + [\boldsymbol{\omega}_\mu, \boldsymbol{\omega}_\nu]. \quad (1.5)$$

Moreover, given a curve in \mathcal{M} , connecting two points x and y , we define the “transport operator”

$$U(x, y) = \mathcal{P} \exp \int_x^y ds \boldsymbol{\omega}_s, \quad (1.6)$$

where $\boldsymbol{\omega}_s$ is the pullback of $\boldsymbol{\omega}_\mu$ on the curve. The path ordering is such that factors corresponding to lower parameters s , that is, closer to x , appear to the left. The matrix $U(x, y) \in \text{SL}(2)$ describes the parallel transport from the local frame at x along the given curve into the local frame at y . A vector $\mathbf{v}(x)$ is mapped onto $\mathbf{v}(y) = U(y, x) \mathbf{v}(x) U(x, y)$, where $U(y, x) = U(x, y)^{-1}$. Note that, at this point, it is important that the tangent bundle of \mathcal{M} is trivial. Otherwise we had to introduce transition functions into the definition (1.6), whenever the curve enters a new coordinate chart.

Lorentz transformations and the time orientation

Under a Lorentz transformations of the local frames, parametrized by an $\text{SL}(2)$ valued field \mathbf{k} , a vector in the local frame at x transforms as $\mathbf{v}(x) \mapsto \mathbf{k}^{-1}(x) \mathbf{v}(x) \mathbf{k}(x)$. For the dreibein and the spin connection, we have

$$\mathbf{e}_\mu \mapsto \mathbf{k}^{-1} \mathbf{e}_\mu \mathbf{k}, \quad \boldsymbol{\omega}_\mu \mapsto \mathbf{k}^{-1} (\partial_\mu + \boldsymbol{\omega}_\mu) \mathbf{k}. \quad (1.7)$$

The metric is invariant, and the transport operator transforms as

$$U(x, y) \mapsto \mathbf{k}^{-1}(x) U(x, y) \mathbf{k}(y). \quad (1.8)$$

Such transformations will be considered as gauge symmetries. They are however not the most general transformations that leave the metric $g_{\mu\nu}$ invariant. A priori, the invariance group of the metric is $\text{O}(1, 2)$. Due to the determinant condition, the symmetry is reduced to $\text{SO}(1, 2)$. However, (1.7) only covers its connected component $\text{SO}_+(1, 2)$ consisting of time oriented proper Lorentz transformations, of which the spinor representation $\text{SL}(2)$ is a two-fold covering.

If we only consider this as the gauge group, then the dreibein contains more physical information than the metric. This extra “bit” of information has to do with the time orientation of spacetime. We can use the dreibein to distinguish between the future and the past of a point in \mathcal{M} . A timelike or lightlike vector v^μ points towards the future, if its image in the local frame $\mathbf{v} = v^\mu \mathbf{e}_\mu$ is causal, that is, positive timelike or lightlike.

This defines a global time orientation of spacetime, which is invariant under gauge transformations of the type (1.7).

The definition of the time orientation of spacetime by the dreibein will be quite useful in the ADM formalism, which we are going to use for the Hamiltonian formulation and quantization of our model. Spacetime is then decomposed into $\mathcal{M} = \mathbb{R} \times \mathcal{N}$, where \mathcal{N} is some two dimensional space manifold, and $t \in \mathbb{R}$ is a time coordinate. This time coordinate should point into the same direction as the physical time. If the surface \mathcal{N} is spacelike, and the dreibein determinant is positive, then this is equivalent to the statement that the normal vector of the ADM surface, which is given by $\varepsilon^{t\mu\nu} e_\mu e_\nu$, is negative timelike. We shall refer to this as the “ADM condition”.

Einstein equations

The Einstein equations in the dreibein formalism can most conveniently be derived from the first order, Palatini type Einstein Hilbert action

$$\frac{1}{16\pi G} \int_{\mathcal{M}} d^3x \varepsilon^{\mu\nu\rho} \text{Tr}(e_\mu \mathbf{F}_{\nu\rho}). \quad (1.9)$$

The dreibein e_μ has the dimension of length, and the field strength $\mathbf{F}_{\mu\nu}$ is dimensionless, so Newtons constant G must have the dimension of an inverse energy, or inverse mass. This explains why there is a natural mass scale in three dimensional gravity.

The vacuum Einstein equations derived from this action, and these are in fact the only ones we need, are the “torsion equation” stating that the dreibein is covariantly constant, and the “curvature equation” requiring the field strength of the spin connection to vanish,

$$\varepsilon^{\mu\nu\rho} D_\nu e_\rho = 0, \quad \varepsilon^{\mu\nu\rho} \mathbf{F}_{\nu\rho} = 0. \quad (1.10)$$

Together they imply that the metric is flat, which means that there are no local gravitational degrees of freedom in the theory. In particular, gravitational waves cannot exist, and there is no gravitational force. Nevertheless, if we insert matter into the universe, there will be effects that very much look like gravitational attraction or scattering.

Embeddings

The general solution to the vacuum Einstein equations is well known [4, 14, 18]. Within a simply connected region $\mathcal{U} \subset \mathcal{M}$, it is parametrized by two scalar fields \mathbf{g} and \mathbf{f} , taking values in the group $\text{SL}(2)$ and the algebra $\mathfrak{sl}(2)$, respectively,

$$\omega_\mu = \mathbf{g}^{-1} \partial_\mu \mathbf{g}, \quad e_\mu = \mathbf{g}^{-1} \partial_\mu \mathbf{f} \mathbf{g}. \quad (1.11)$$

The metric and the dreibein determinant in \mathcal{U} becomes

$$g_{\mu\nu} = \frac{1}{2} \text{Tr}(\partial_\mu \mathbf{f} \partial_\nu \mathbf{f}), \quad e = \frac{1}{12} \varepsilon^{\mu\nu\rho} \text{Tr}(\partial_\mu \mathbf{f} \partial_\nu \mathbf{f} \partial_\rho \mathbf{f}), \quad (1.12)$$

and the transport operator for a curve that entirely lies inside \mathcal{U} is given by

$$\mathbf{U}(x, y) = \mathbf{g}^{-1}(x) \mathbf{g}(y). \quad (1.13)$$

From these relations we infer the following physical interpretation for g and f . The map f provides an isometric embedding of \mathcal{U} into Minkowski space. The metric on \mathcal{U} is equal to the metric induced by the embedding, and f is locally one-to-one, provided that the determinant is positive, which is equal to the Jacobian of f . Hence, every simply connected region of spacetime looks like a piece of Minkowski space.

There is also an interpretation for g . Consider the expression for the transport operator above. The parallel transport is independent of the path, and it is carried out in two steps. A vector (or spinor) in the local frame at x is first mapped into a “background frame” by $g^{-1}(x)$, and from this background frame it is then mapped into the local frame at y by $g(y)$. The field g identifies all the local frames inside \mathcal{U} with a single background frame, which is the one associated with the embedding Minkowski space.

So, we can say that the pair (g, f) provides an embedding of the complete bundle of local frames over \mathcal{U} into Minkowski space. Every quantity, for example a vector $v(x)$ in the local frame at x , can be equivalently represented in the background frame as $w = g(x)v(x)g^{-1}(x)$. To see that this is consistent, consider a local Lorentz transformation (1.7). Applying this to (1.11), we find that

$$g \mapsto gk, \quad f \mapsto f. \quad (1.14)$$

Hence, the embedding of spacetime into Minkowski space is not affected. The vector w in the background frame is also unchanged, because under a local Lorentz transformation we have $v(x) \mapsto k^{-1}(x)v(x)k(x)$. The background space is invariant under local Lorentz transformations. But there is another symmetry group that naturally acts on it. The right hand sides of (1.11) are invariant under rigid Poincaré transformations,

$$g \mapsto h^{-1}g, \quad f \mapsto h^{-1}(f - n)h, \quad (1.15)$$

where $h \in \text{SL}(2)$ and $n \in \mathfrak{sl}(2)$ are *constants*. An isometric embedding of a piece of flat spacetime into Minkowski space is only determined up to such a transformation. If we think of the background space as a rest frame of some, say, external observer, then it is the location and the orientation of this observer that is changed by (1.15).

The freedom to perform a rigid Poincaré transformation shows that the background frame, or the observer associated with it, has no particular physical meaning. It is just an auxiliary construction, which was used to write down the solutions to the vacuum Einstein equations. If we parametrize these solutions by g and f , then the Poincaré transformations (1.15) should be considered as gauge symmetries. This will change when we derive the action for the particle model in section 3. In order to obtain a well defined action, we have to add an observer to the universe, which will essentially be described by such an embedding, and this observer will become a physical reality. That is, the action will explicitly depend on the fields g and f , so that the background frame is no longer a gauge degree of freedom.

2 The particle

Let us now turn to the description of a point particle. In three dimensional Einstein gravity, spacetime in the neighbourhood of a massive, pointlike particle has the shape of

a cone, with the particle sitting at the top. As already mentioned in the introduction, the deficit angle α of the cone is related to the mass m of the particle by $\alpha = 8\pi G m$. To obtain a dimensionless parameter m , and to avoid unnecessary factors in the formulas, we set $G = 1/4\pi$ in the following, so that $\alpha = 2m$. In the neighbourhood of the world line of such a particle, we can introduce cylindrical coordinates (t, r, φ) , such that the metric becomes

$$ds^2 = -dt^2 + dr^2 + \left(1 - \frac{m}{\pi}\right)^2 r^2 d\varphi^2. \quad (2.1)$$

Here, we have $r \geq 0$, and the angular coordinate φ is periodic with a period of 2π . The total angle of this conical singularity is $2\pi - 2m$, which is the ratio of the circumference and the radius of a small circle centered at $r = 0$. So, the deficit angle is indeed equal to $2m$.

There are some special values for the mass which restrict the “physically reasonable” range of the parameter m . First of all, the deficit angle of a cone cannot be larger than 2π , and thus the mass should not exceed π . In fact, we can restrict the parameter m in (2.1) to be less than or equal to π , without loss of generality, because replacing m with $2\pi - m$ does not affect the metric. On the other hand, the mass of the particle should also be non-negative. The allowed range for the parameter m is therefore $0 \leq m \leq \pi$.

To see that negative masses lead to “unphysical” situations, consider the metric (2.1) for $m < 0$. In this case, the total angle of the conical singularity is larger than 2π . This implies that there is negative curvature, and thus a negative energy density at the top of the cone. As a consequence, the particle is gravitationally repulsive. Two parallel geodesics, interpreted as test particles, passing the particle on opposite sides, are bent away from each other after they have passed the singularity. In other words, they are repelled from the matter source.

Another question is whether we should include the bounds 0 and π into the range of m or not. If we consider the metric given above, then both cases are somewhat problematic. For $m = 0$, we recover flat Minkowski space, and the particle disappears completely. For $m = \pi$, the metric becomes singular and the cone degenerates to a line. However, in both cases, this is only due to the special coordinate system. It describes the particle in its own rest frame. We should expect that, for $m = 0$, the world line of the particle becomes lightlike, so that a rest frame no longer exist. This will in fact be the case, and, quite surprisingly, the same will happen at the upper bound $m = \pi$.

Regularization

For a proper description of a pointlike particle within the framework of matter coupled Einstein gravity as a field theory, we somehow have to get rid of the curvature singularity at $r = 0$. There are basically two known ways to deal with this problem. The first possibility is to introduce a *puncture*. That is, one excludes the world line of the particle from the spacetime manifold. The metric is then regular everywhere on the remaining manifold, and the particle is described by an appropriate set of boundary conditions imposed on the fields in the neighbourhood of the puncture. The second possibility is to *regularize* the matter source, for example by replacing the pointlike

particle by some, sufficiently smooth, cylindrical matter source centered around the world line. Then one does not need to impose boundary conditions, but instead one has to deal with more complicated, matter coupled Einstein equations, and in the end one has to take a limit where the radius of the cylinder shrinks to zero.

Here we shall use a new method which is somehow a mixture of these, and which makes as much as possible use of the advantages of both methods. It has in common with the puncture method that it essentially converts the matter degrees of freedom carried by the particle into topological degrees of freedom of the gravitational field, associated with a non-contractible loop that surrounds the puncture. On the other hand, it also works somewhat like the second method. The world line of the particle is smeared out, in a way which is sufficient to remove the singularity. It is however not necessary to take any kind of limit to remove the regularization. In particular, the metric in the neighbourhood of the world line will not be affected by the regularization.

The basic idea is quite simple. Consider the covariant metric given by (2.1). As it stands, it is not immediately obvious that there is a singularity. It is given as a smooth function of the coordinates (t, r, φ) and their differentials. The *curvature* singularity cannot be seen because there is also a *coordinate* singularity at $r = 0$. To get rid of the curvature singularity, it might be sufficient to remove the coordinate singularity. This can be done, by changing the topology of spacetime. Instead of interpreting the cylindrical coordinates in the usual way, we do not identify the points with different φ at $r = 0$. As a result, spacetime is no longer an \mathbb{R}^3 , but a manifold with a *cylindrical boundary*. This boundary represents the world line of the particle, which is thus, in a sense, smeared out. But there is also a non-contractible loop surrounding the particle, so that topological degrees of freedom of the gravitational field can arise.

On the modified spacetime manifold, it is now rather obvious that the metric becomes a smooth second rank tensor. The coordinate system (t, r, φ) is everywhere regular. The pullback of the metric on the boundary, where (t, φ) are the standard coordinates on a cylinder, is given by $ds^2 = -dt^2$. So, the only problem is that the metric is singular on the boundary. It is of rank one instead of two, which is the dimension of the boundary. This reflects the fact that the *two dimensional* cylindrical boundary actually looks like a *one dimensional* world line of a particle. In the dreibein formulation of gravity, exactly these kinds of singular metrics can be consistently included. A possible dreibein representation of (2.1) is given by

$$\begin{aligned} e_t &= \gamma_0, & e_r &= \cos \varphi \gamma_1 + \sin \varphi \gamma_2, \\ e_\varphi &= \left(1 - \frac{m}{\pi}\right) r (\cos \varphi \gamma_2 - \sin \varphi \gamma_1). \end{aligned} \quad (2.2)$$

It is chosen such that the determinant is positive, the ADM condition holds (with $\varepsilon^{tr\varphi} = 1$), and the spin connection takes the simplest possible form,

$$\omega_t = 0, \quad \omega_r = 0, \quad \omega_\varphi = -\frac{m}{2\pi} \gamma_0. \quad (2.3)$$

Obviously, all these fields are smooth, and they provide a solution to the *vacuum* Einstein equations on the modified spacetime manifold. The curvature singularity is completely removed, although the (smeared) world line of the particle is still included in the spacetime manifold.

We can summarize this as follows. A suitable way to include point particles into the dreibein formulation of gravity is to consider their world lines as cylindrical boundaries of spacetime. As we are, within this article, only interested in a universe containing exactly one such particle, we can fix the spacetime manifold to be $\mathcal{M} = \mathbb{R} \times \mathbb{R}_+ \times S^1$. The cylindrical coordinates (t, r, φ) then form a global coordinate system on \mathcal{M} . It is also easy to see that \mathcal{M} satisfies the assumption made in the previous section, namely, that its tangent bundle is trivial. The only transition function, arising due to the identification $\varphi \equiv \varphi + 2\pi$, is the identity.

The only necessary modification to the definition of dreibein gravity in the previous section is that we have to weaken the determinant and ADM conditions slightly. To allow singular metrics on the boundary, we have to require $e > 0$ in the *interior* of \mathcal{N} only. The metric can then be of a lower rank on the boundary, and it will in fact be forced to be of rank one by the field equations. The same also holds for the ADM condition, saying that the normal vector of the hypersurfaces of constant time coordinate should be negative timelike. For (2.2), this vector vanishes on the boundary as well, so that the ADM condition also applies in the interior of \mathcal{N} only.

The point particle condition

The vacuum Einstein equations are not sufficient to make the boundary of \mathcal{M} look like a world line of a point particle with a given mass. We need some additional field equations on the boundary. One of them has to ensure that the metric on the boundary is of rank one, so that it looks like a line. We shall refer to this as the “point particle condition”. Let us first deal with this, and then return to the mass of the particle later on. If we want a cylinder to look like a line, then its circumference has to vanish. This means that all φ -components of the covariant metric have to vanish at $r = 0$, or, equivalently,

$$\bar{e}_\varphi = 0. \quad (2.4)$$

Here, the bar is used to denote the value of a field on the boundary. Obviously, the dreibein (2.2) satisfies this condition. But what is the general solution to (2.4) and the Einstein equations? For the vacuum Einstein equations, we know this already. Inside a simply connected region, it is given by an embedding (g, f) of spacetime into Minkowski space. Now, the neighbourhood of the particle is not simply connected, but we can cover it by a simply connected region. If we introduce a cut along, say, the plane at $\varphi = \pm\pi$ (see figure 2), then we can write the general solution to the Einstein equations in the neighbourhood of the boundary as in (1.11),

$$\omega_\mu = g^{-1} \partial_\mu g, \quad e_\mu = g^{-1} \partial_\mu f g, \quad (2.5)$$

where g and f are defined on the range $-\pi \leq \varphi \leq \pi$, and, of course, they also depend on t and r . Their values at $\varphi = \pm\pi$, denoted by g_\pm and f_\pm , are not independent. They need not be equal, but the one-forms e_μ and ω_μ must be continuous, which means that the right hand sides of (2.5) at $\varphi = -\pi$ must coincide with those at $\varphi = \pi$. We saw that this is the case if and only if the fields g and f are related by a Poincaré transformation (1.15). Hence, we must have

$$g_+ = u^{-1} g_-, \quad f_+ = u^{-1} (f_- - v) u, \quad (2.6)$$

where $\mathbf{u} \in \text{SL}(2)$ and $\mathbf{v} \in \mathfrak{sl}(2)$ are constants. In addition to this, we also have to take into account the point particle condition. Expressed in term of \mathbf{g} and \mathbf{f} , it becomes a simple differential equation,

$$\bar{e}_\varphi = \bar{\mathbf{g}}^{-1} \partial_\varphi \bar{\mathbf{f}} \bar{\mathbf{g}} = 0, \quad (2.7)$$

which states that the value of \mathbf{f} on the boundary must be a function of t only, $\bar{\mathbf{f}}(t, \varphi) = \mathbf{x}(t)$. This function is further restricted by the condition (2.6), which yields

$$\mathbf{v} = \mathbf{x}(t) - \mathbf{u} \mathbf{x}(t) \mathbf{u}^{-1}, \quad (2.8)$$

because $\bar{\mathbf{f}}_+(t) = \bar{\mathbf{f}}_-(t) = \mathbf{x}(t)$. Now, \mathbf{u} and \mathbf{v} are constants, and therefore the time derivative of this equation implies that the time derivative of \mathbf{x} must commute with \mathbf{u} . All vectors that commute with a given group element $\mathbf{u} \in \text{SL}(2)$ are proportional to its “projection” $\mathbf{p} \in \mathfrak{sl}(2)$ into the algebra (see (A.9)). This vector \mathbf{p} is also the “axis” of the Lorentz transformation represented by \mathbf{u} . It can be obtained by expanding the matrix \mathbf{u} in terms of the unit and gamma matrices, and then dropping the contribution proportional to the unit matrix,

$$\mathbf{u} = u \mathbf{1} + p_a \gamma^a \quad \mapsto \quad \mathbf{p} = p_a \gamma^a. \quad (2.9)$$

Taking all this together, we find that the most general function \mathbf{f} is, on the boundary, of the form

$$\bar{\mathbf{f}}(t, \varphi) = \mathbf{y} + \tau(t) \mathbf{p}, \quad (2.10)$$

where $\mathbf{y} \in \mathfrak{sl}(2)$ is some constant and τ is an arbitrary real function of time. Obviously, this reminds of the description of a relativistic point particle in Minkowski space. The particle has a momentum \mathbf{p} , and it passes through the point \mathbf{y} . Indeed, with the interpretation of \mathbf{f} as an embedding of spacetime into Minkowski space, $\bar{\mathbf{f}}$ is the position of the particle in this background space. If it is considered as the rest frame of an external observer, then he sees the particle moving along the world line (2.10). The arbitrary function $\tau(t)$ has physical interpretation as well. Consider the time component of the dreibein on the boundary,

$$\bar{e}_t = \bar{\mathbf{g}}^{-1} \partial_t \bar{\mathbf{f}} \bar{\mathbf{g}} = \partial_t \tau \bar{\mathbf{g}}^{-1} \mathbf{p} \bar{\mathbf{g}}. \quad (2.11)$$

It represents the tangent vector of the world line, or the velocity of the particle, in the local frames on the boundary. Its length is given by

$$\sqrt{-\bar{g}_{tt}} = |\partial_t \tau| \sqrt{-p_a p^a}, \quad (2.12)$$

and this is the quotient of the physical time elapsing on the world line divided by the coordinate time. The function τ is thus, up to a factor, the eigentime of the particle.

The mass shell condition

To see that the vector \mathbf{p} , or rather the group element \mathbf{u} from which it is derived, behaves very much like the momentum of a relativistic point particle, let us now, before we

come back to the solution of the field equations in the neighbourhood of the particle, consider the second boundary condition to be imposed. So far, we haven't fixed the mass of the particle. It should be related to the deficit angle by $\alpha = 2m$, and, like the point particle condition, we want to write this “mass shell condition” as a local field equation, that is, it should involve the fields on the boundary only.

At first sight, it is not so obvious how this can be achieved, because the deficit angle, defined as the missing circumference of a small circle, is a property of the metric *in the neighbourhood* of the particle. However, if we slightly change the definition of the mass, we can express it as a function of the spin connection *on the boundary*. The idea is that, instead of measuring the radius and circumference of a circle, we can also perform a parallel transport of a vector (or spinor) around the particle. If the deficit angle is $2m$, then the result should be that the vector is rotated by $2m$.

The parallel transport along a given curve is described by the transport operator (1.6). Provided that the curvature vanishes, this is invariant under smooth deformations of the curve. If we are interested in the effect of transporting a vector or spinor around the particle, we can, in particular, perform this transport on the boundary. Let us therefore define an $\text{SL}(2)$ valued field $\bar{\mathbf{u}}(t, \varphi)$, which describes the transport of a spinor once around the boundary, on a circle of constant t , starting and ending at the point φ . For later convenience, we choose the “clockwise”, that is, negative φ direction. The “holonomy” of the particle is then given by

$$\bar{\mathbf{u}}^{-1}(t, \varphi) = \mathcal{P} \exp \int_{\varphi}^{\varphi+2\pi} d\tilde{\varphi} \bar{\omega}_{\varphi}(t, \tilde{\varphi}). \quad (2.13)$$

Independently of whether the spin connection is flat or not, this is a well defined field on the boundary. It is, by definition, covariantly constant with respect to φ ,

$$D_{\varphi} \bar{\mathbf{u}} = \partial_{\varphi} \bar{\mathbf{u}} + [\bar{\omega}_{\varphi}, \bar{\mathbf{u}}] = 0, \quad (2.14)$$

but there is, a priori, no relation between the values of $\bar{\mathbf{u}}$ at different times t . We can therefore think of $\bar{\mathbf{u}}$ as a function on the world line of the particle, which “essentially” depends on t only. The dependence of $\bar{\mathbf{u}}$ on φ is only due to the fact that the same physical quantity is represented in different local frames on the boundary.

If the Einstein equations are satisfied, then $\bar{\mathbf{u}}$ is covariantly constant with respect to t as well. The physical quantity represented by $\bar{\mathbf{u}}$ is thus a constant of motion. It is in fact the same as the constant \mathbf{u} introduced above, which we interpreted as the *momentum* of the particle. The relation between $\bar{\mathbf{u}}$ and \mathbf{u} , or that between their projections $\bar{\mathbf{p}}$ and $\bar{\mathbf{p}}$, is found to be

$$\bar{\mathbf{u}} = \bar{\mathbf{g}}^{-1} \mathbf{u} \bar{\mathbf{g}}, \quad \bar{\mathbf{p}} = \bar{\mathbf{g}}^{-1} \mathbf{p} \bar{\mathbf{g}}. \quad (2.15)$$

The right hand sides are continuous at $\varphi = \pm\pi$, as a consequence of (2.6), although $\bar{\mathbf{g}}$ is not. These relations are also the reason for choosing the clockwise direction in the definition of $\bar{\mathbf{u}}$. Otherwise, there would be an “inverse” showing up in (2.15).

Now, coming back to the mass shell condition, we have to specify all holonomies $\bar{\mathbf{u}}$ which represent rotations by $2m$. This is a simple geometric exercise, which is

explicitly carried out in Appendix A. The result is that all such elements of $\text{SL}(2)$ can be characterized by the equation (A.18),

$$\frac{1}{2}\text{Tr}(\bar{\mathbf{u}}) = \cos m, \quad (2.16)$$

The solutions to this equation lie in the conjugacy classes of the group elements $e^{\pm m\gamma_0}$. Geometrically, the sign in the exponent distinguishes between rotations in positive and negative direction. In the context of particle physics, there is also another interpretation of the sign ambiguity. Let us write this “mass shell condition” as a condition to be imposed on the momentum vector \mathbf{p} . We first notice that the momentum \mathbf{u} in the background frame satisfies the same mass shell condition as $\bar{\mathbf{u}} = \bar{\mathbf{g}}^{-1}\mathbf{u}\bar{\mathbf{g}}$, because the conjugation with $\bar{\mathbf{g}}$ drops out under the trace. Furthermore, if we use the expansion (2.9), and that the vector p_a is related to the scalar u by

$$u^2 = p_a p^a + 1, \quad (2.17)$$

then we find

$$u = \cos m \quad \Rightarrow \quad p_a p^a + \sin^2 m = 0. \quad (2.18)$$

This looks like the mass shell condition of a relativistic point particle, except that the mass is replaced by its sine. This is, in the first place, a matter of convention, and follows from the fact that we defined the momentum vector \mathbf{p} to be the *projection* of \mathbf{u} into the algebra. One can also define it such that \mathbf{u} is the *exponential* of \mathbf{p} [6]. In this case, there would be no sine in (2.18), but we would run into trouble later on, because the exponential map is neither one-to-one, nor does it map the algebra *onto* the group, and so \mathbf{p} would not be a well defined phase space function. In any case, the structure of the solutions to (2.18) is always the same. They split into two subsets, the “particle” and the “antiparticle” mass shell, consisting of the positive and negative timelike vectors of length $\sin m$.

To visualize these mass shells, let us use the coordinates (p_a, u) to embed the group manifold into \mathbb{R}^4 . The condition (2.17) defines a hyperboloid therein, which is shown in figure 1. The mass shells are the intersections of this hyperboloid with the plane $u = \cos m$. We see that there is an upper and a lower mass shell, corresponding the positive and negative timelike vectors \mathbf{p} , and that they look very similar to those of a relativistic point particle. But there are also some features that are different.

One essential difference is that the range of m is bounded from below *and* from above. We can now see that this is because the momentum \mathbf{u} lives on the group manifold $\text{SL}(2)$, and not in flat Minkowski space. If the mass approaches the lower bound $m = 0$, then the mass shells approach the “light cones” emerging from $\mathbf{u} = 1$. They consist of those elements of the group for which $u = 1$, so that \mathbf{p} is a lightlike vector. But now there is a second pair of light cones, emerging from the group element $\mathbf{u} = -1$. There, we have $u = -1$, and so \mathbf{p} is also lightlike. For $m = \pi$, the mass shells coincide with these light cones.

Between these two pairs of light cones, there is only a finite range of u , which coincides with the range of the cosine in the mass shell condition. The whole range of timelike momenta \mathbf{u} is covered by $0 < m < \pi$, and on both sides of this interval the

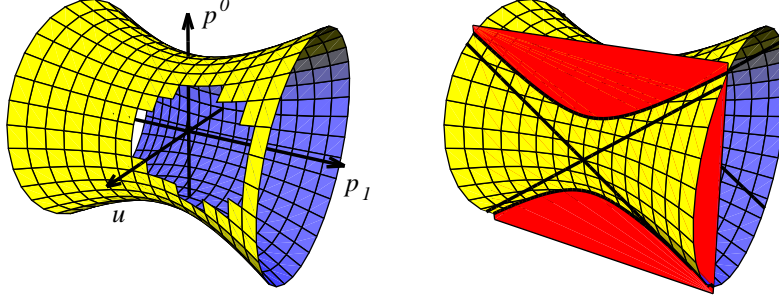


Figure 1: The group manifold $SL(2)$, embedded in \mathbb{R}^4 , using the coordinates $p_A = (p_a, u)$, with p_2 suppressed. The picture to the right shows the two mass shells for $m = \pi/6$ (a deficit angle of 60°), obtained by intersecting the group manifold with the plane $u = \cos m$, and the light cones emerging from $\mathbf{1}$ and $-\mathbf{1}$ (on the back). The grid lines on the group manifold are the Euler angles ρ and χ .

momentum becomes lightlike. Using the conventional terminology, we can say that the particle is “massless” for $m = 0$ as well as for $m = \pi$. This is what we already mentioned in the beginning. It is the reason why the description of the particle in its own rest frame fails for these values of m .

To see what goes wrong if we take the limit $m \rightarrow 0$ or $m \rightarrow \pi$ in the rest frame of the particle, consider the dreibein (2.2) and the spin connection (2.3). The momentum is then given by

$$\bar{\mathbf{u}} = e^{m\gamma_0} = \cos m \, \mathbf{1} + \sin m \, \gamma_0. \quad (2.19)$$

In the limits $m \rightarrow 0$ and $m \rightarrow \pi$, we have $\bar{\mathbf{u}} \rightarrow \pm \mathbf{1}$ and therefore $\bar{\mathbf{p}} \rightarrow 0$. The same happens for a relativistic point particle if we take the limit $m \rightarrow 0$ in the rest frame. The momentum does not end up on the light cone, but vanishes. To get a proper description of lightlike particles, we have to exclude the special solutions $\bar{\mathbf{u}} = \pm \mathbf{1}$ of the mass shell condition. The neighbourhood of a lightlike particle will then, as we shall see in a moment, not be very different from that of a massive particle.

The neighbourhood of the particle

Let us now return to the general solution of the Einstein equations in the neighbourhood of the particle. So far, we have solved the point particle condition and the Einstein equations *on the boundary*. To see what spacetime looks like in the neighbourhood, we have to find an embedding (g, f) , such that it satisfies the condition (2.6) along the cut at $\varphi = \pm\pi$,

$$g_+ = u^{-1} g_-, \quad f_+ = u^{-1} (f_- - v) u, \quad (2.20)$$

and such that, on the boundary, \mathbf{f} takes the form (2.10),

$$\bar{\mathbf{f}}(t, \varphi) = \mathbf{y} + \tau(t) \mathbf{p}. \quad (2.21)$$

Moreover, it is subject to some additional conditions. The resulting metric must be invertible, and the surfaces of constant t must be spacelike. We can further restrict it by imposing “gauge conditions”. In particular, we are still free to choose the coordinates (t, r, φ) off the boundary as we like, and we can use this to make the function \mathbf{f} look as simple as possible.

We found that the determinant condition for the dreibein is equivalent to the statement that \mathbf{f} is locally one-to-one. We can therefore use \mathbf{f} to *define* the coordinates. Let us fix the time coordinate t such that the surfaces of constant t in \mathcal{M} are mapped onto equal time planes in the background Minkowski space. By this, we mean the planes of constant $f_0 = \frac{1}{2} \text{Tr}(\mathbf{f} \gamma_0)$. The value of the zero component of \mathbf{f} is then determined by its value on the boundary.

Furthermore, we can choose the radial coordinate r to be the physical distance from the particle. The difference of \mathbf{f} to its value on the boundary is then a spacelike vector of length r , lying in the γ_1 - γ_2 plane. The only remaining freedom is the direction of this vector, which can be specified by a real function $\theta(t, r, \varphi)$. All together, the most general function \mathbf{f} with all these properties is of the form

$$\mathbf{f}(t, r, \varphi) = \mathbf{y} + \tau \mathbf{p} + r e^{\frac{1}{2}\theta \gamma_0} \gamma_1 e^{-\frac{1}{2}\theta \gamma_0}, \quad (2.22)$$

or, in components,

$$\begin{aligned} f_0 &= y_0 + \tau p_0, & f_1 &= y_1 + \tau p_1 + r \cos \theta, \\ f_2 &= y_2 + \tau p_2 + r \sin \theta. \end{aligned} \quad (2.23)$$

The situation is sketched in figure 2. The neighbourhood of the boundary, which is cut along the surface at $\varphi = \pm\pi$, is mapped onto a “wedge” in Minkowski space. Thereby, the ADM surface shown on the left is mapped onto the equal time surface on the right. The images of the cut at $\varphi = \pm\pi$ are marked θ_{\pm} . They will turn out to be planes, as a consequence of the boundary conditions for \mathbf{f} , but for the moment they are arbitrary surfaces described by two functions,

$$\theta_+(t, r) = \theta(t, r, \pi), \quad \theta_-(t, r) = \theta(t, r, -\pi). \quad (2.24)$$

Let us now check the ADM and the determinant condition. The normal vector of the constant time surface is found to be

$$\frac{1}{2} \varepsilon^{t\mu\nu} e_\mu e_\nu = e_{[r} e_{\varphi]} = \mathbf{g}^{-1} \partial_{[r} \mathbf{f} \partial_{\varphi]} \mathbf{f} \mathbf{g} = -r \partial_\varphi \theta \mathbf{g}^{-1} \gamma_0 \mathbf{g}, \quad (2.25)$$

and the dreibein determinant is

$$e = \frac{1}{2} \text{Tr}(e_t e_r e_\varphi) = \frac{1}{2} \text{Tr}(\partial_t \mathbf{f} \partial_r \mathbf{f} \partial_\varphi \mathbf{f}) = r p^0 \partial_t \tau \partial_\varphi \theta. \quad (2.26)$$

For $r > 0$, the normal vector must be negative timelike and the determinant has to be positive. This is the case if and only if

$$\partial_\varphi \theta > 0, \quad p^0 \partial_t \tau > 0. \quad (2.27)$$

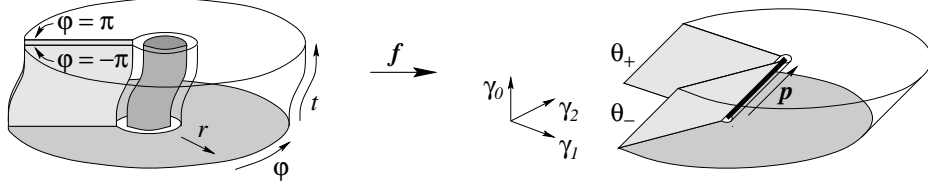


Figure 2: The embedding of a neighbourhood of the cylindrical boundary of spacetime (to the left) into the background Minkowski space (to the right). Thereby, the boundary at $r = 0$ is mapped onto the world line pointing into the direction of \mathbf{p} , the two sides of the cut at $\varphi = \pm\pi$ are mapped onto the surfaces θ_{\pm} , and the ADM surface of constant t is mapped onto a surface of constant f_0 .

The first inequality says that θ must be an increasing function of φ , which means that the embedding shown in figure 2 is such that, with increasing φ , we move counter-clockwise around the world line of the particle in the embedding space. In particular, we must have $\theta_+ > \theta_-$. The second one implies that τ is an increasing or decreasing function of time, depending on whether p^0 is positive or negative. It cannot be fulfilled if p^0 is zero, but then the momentum vector would be either spacelike or zero, which is excluded by the mass shell condition.

The relation between t , τ and the sign of p^0 confirms very nicely the interpretation of the states with positive and negative p^0 as particle and antiparticle states. Particles have positive timelike or lightlike momentum vectors, and their eigentime τ is increasing with the external time t . For antiparticles, the momentum vector is negative timelike or lightlike, and the eigentime is running backwards. This is another analogy to the relativistic point particle.

Now, we still have to check the boundary condition (2.20). For this purpose, it is most convenient to express the holonomy $\mathbf{u} \in \text{SL}(2)$ in terms of “Euler angles” (see (A.11)). They provide a global set of coordinates on the group manifold $\text{SL}(2)$. A generic group element can be written as

$$\begin{aligned} \mathbf{u} &= e^{\frac{1}{2}(\rho+\phi)\gamma_0} e^{\chi\gamma_1} e^{\frac{1}{2}(\rho-\phi)\gamma_0} \\ &= \cosh \chi (\cos \rho \mathbf{1} + \sin \rho \gamma_0) + \sinh \chi (\cos \phi \gamma_1 + \sin \phi \gamma_2). \end{aligned} \quad (2.28)$$

Here, ρ and ϕ are periodic “angular” coordinates with period 2π , and χ is a positive “radial” coordinate, which forms together with ϕ a polar coordinate system (for $\chi = 0$ the angle ϕ is redundant). The coordinates ρ and χ are shown in figure 1, as the grid lines. The angular coordinate ρ is the one that winds around the hyperboloid, and χ increases with the distance from the center. We can also express the momentum vector \mathbf{p} in terms of these variables. It becomes

$$\mathbf{p} = \cosh \chi \sin \rho \gamma_0 + \sinh \chi (\cos \phi \gamma_1 + \sin \phi \gamma_2), \quad (2.29)$$

and this tells us that the particle is moving with a velocity of $\tanh \chi / \sin \rho$ (which is 1 for lightlike and smaller than one for timelike vectors \mathbf{p}) into the spatial direction specified by the angle ϕ . We can distinguish particles and antiparticles using the coordinate

ρ . We have a particle state for $0 < \rho < \pi$, and an antiparticle state for $-\pi < \rho < 0$. As ρ has a period of 2π and cannot be a multiple of π (in this case p^0 would be zero), it always lies in one of these intervals.

Inserting all this into the boundary condition (2.20) for \mathbf{f} , we obtain a relation between the functions θ_{\pm} and the Euler angles, which after some algebraic manipulations becomes

$$e^{-\frac{1}{2}(\theta_+ + \rho + \phi)\gamma_0} e^{-\chi\gamma_1} e^{\frac{1}{2}(\theta_- - \rho + \phi)\gamma_0} \gamma_1 e^{-\frac{1}{2}(\theta_- - \rho - \phi)\gamma_0} e^{\chi\gamma_1} e^{\frac{1}{2}(\theta_+ + \rho - \phi)\gamma_0} = \gamma_1 \quad (2.30)$$

Let us first consider the case where $\chi = 0$, describing a particle which is at rest in the embedding Minkowski space. In this case, (2.30) reduces to

$$e^{-\frac{1}{2}(\theta_+ - \theta_- + 2\rho)\gamma_0} \gamma_1 e^{\frac{1}{2}(\theta_+ - \theta_- + 2\rho)\gamma_0} = \gamma_1. \quad (2.31)$$

According to the formula (A.13), this implies that the term between the parenthesis must be a multiple of 2π , so that

$$\theta_+ - \theta_- = 2\pi z - 2\rho, \quad z \in \mathbb{Z}. \quad (2.32)$$

In particular, the difference $\theta_+ - \theta_-$ must be a constant. This is the total angle of the cone surrounding the particle. It is determined by the momentum of the particle up to a multiple of 2π . The same holds for the deficit angle, which is given by $\alpha = 2\pi - \theta_+ + \theta_-$. The integer z is however not completely arbitrary. For $\theta_+ > \theta_-$, we must have $z \geq 1$ for particles and $z \geq 0$ for antiparticles. All together, we find that the possible values for the deficit angle are given by

$$\alpha = 2\rho, \quad 2\rho - 2\pi, \quad 2\rho - 4\pi, \quad \dots \quad (2.33)$$

for a particle state with $0 < \rho < \pi$, and

$$\alpha = 2\rho + 2\pi, \quad 2\rho, \quad 2\rho - 2\pi, \quad \dots \quad (2.34)$$

for an antiparticle state with $-\pi < \rho < 0$. Obviously, only the first elements of these series are positive. In all other cases, the deficit angle is negative. The total angle is then bigger than 2π , which means that the wedge shown in figure 2 overlaps itself. In the beginning, we argued that such particles are “unphysical”, because they are gravitationally repulsive. To exclude them, we can impose yet another boundary condition on the metric.

It is not necessary to specify this condition explicitly as, say, a function of the metric in the neighbourhood of the particle. It is sufficient to know that, for any given momentum \mathbf{u} , there is a discrete series of solutions to the Einstein equations in the neighbourhood of the particle, and that the condition $\alpha > 0$ picks out exactly one of these solutions. To be precise, so far we have only shown this for a particle at rest, with $\chi = 0$. For $\chi > 0$, the condition (2.30) is more restrictive. Evaluating it using the formulas (A.12), one finds that

$$\theta_+ = \phi - \rho + 2\pi z_+, \quad \theta_- = \phi + \rho - 2\pi z_-, \quad (2.35)$$

where z_{\pm} are either two integers or two half integers. In particular, the right hand sides of these equations are constants, and thus the surfaces θ_{\pm} must be planes. Subtracting the two equations, we get (2.32) back. The relation between ρ and the deficit angle is therefore also valid for moving particles.

The additional information contained in (2.35) is the following. It is not only the total angle that is determined by the momentum, but also the directions θ_+ and θ_- themselves are fixed. If we add the two equations, then we find that the angle between the directions θ_+ and ϕ must be the same as that between ϕ and θ_- . As the Euler angle ϕ was the direction in which the particle is moving, this means that the two images of the cut in the embedding Minkowski space have to lie symmetrically “in front of” or “behind” the particle. The latter situation is shown in the figure.

This reproduces very nicely an observation made by 't Hooft [9], who obtained the same result in a somewhat different way. If one wants the ADM surfaces to become equal time planes in an observer's rest frame, which is not the rest frame of the particle, then the wedge that the particle “cuts out” from Minkowski space has to be placed either straight in front of or behind the particle. We should however emphasize that this is just a result of the special gauge choice, that is, the special choice of coordinates in the neighbourhood of the particle. This was only made to show what this neighbourhood looks like. We shall not impose any such gauge fixing on our model.

The rest mass and the energy

We have now found a relation between the momentum u of the particle and the deficit angle of the conical singularity on its world line. But this seems to be somewhat in contradiction with the previously considered relation $\alpha = 2m$. Instead of this, we now have $\alpha = 2\rho$ for particles with $0 < \rho < \pi$, and $\alpha = 2\pi + 2\rho$ for antiparticles with $-\pi < \rho < 0$. Furthermore, the deficit angle is *frame dependent*. It depends on the relative motion of the particle with respect to the observer. To find a relation between the deficit angle and the mass, we have to express the mass shell condition in terms of the Euler angles. It becomes

$$\cosh \chi \cos \rho = \cos m. \quad (2.36)$$

It provides a relation between the mass m , the “rapidity” χ , which measures the velocity of the particle with respect to the observer, and the deficit angle. If the particle is at rest, then χ vanishes and this reduces to $\cos \rho = \cos m$. There are two solutions to this equation, $\rho = \pm m$. For the positive sign, we have a particle state, and we recover the original relation $\alpha = 2m$. If we choose the negative sign, then we have an antiparticle state, and the deficit angle becomes $\alpha = 2\pi - 2m$.

Actually, one should expect the gravitational field of an antiparticle to be the same as that of the corresponding particle. But, in a somewhat heuristical way, it can be understood as a result of the unusual spectrum of the parameter m . Antiparticles are related to particles by a charge conjugation operation, which inverts all charges, including the gravitational charge, and thus the mass. But the only reasonable way to “invert” the mass, which takes values in the range $0 \leq m \leq \pi$, is to replace it by $\pi - m$.

Apart from this somewhat strange effect, a more interesting question is, what happens if we leave the rest frame of the particle. The deficit angle is then no longer equal

to $2m$ or $2\pi - 2m$. Geometrically, this can be understood as follows. If the particle is accelerated, then the wedge in figure 2 is boosted, so that the angle between θ_+ and θ_- changes. This explains why the deficit angle is a dynamical quantity. Let us see how it behaves for small masses and small velocities. We can then expand the mass shell condition (2.36), and what we get is

$$(1 + \frac{1}{2}\chi^2)(1 - \frac{1}{2}\rho^2) \approx (1 - \frac{1}{2}m^2) \Rightarrow \rho^2 \approx \chi^2 + m^2. \quad (2.37)$$

Now, if χ is small, then it is approximately the length of the spatial component of the momentum (2.29), and the above relation becomes the usual relation between the mass, the spatial momentum, and the energy of a relativistic point particle. This suggests that we should interpret the deficit angle as the *energy* of the particle rather than its *mass*. The relation (2.36) is then a somewhat deformed dispersion relation, which reduces to the usual one in the low energy limit.

At large scales, the behaviour of ρ as a function of χ is however completely different. For large χ , ρ always saturates at $\pi/2$ for particles and at $-\pi/2$ for antiparticles. Whether it approaches this limit from below or from above depends on whether its rest mass is smaller or bigger than $\pi/2$. This can also be seen in figure 1, where the grid lines on the hyperboloid represent the Euler angles. For large momenta, the value of ρ on the upper mass shell always approaches that on the “ridge” of the hyperboloid, which is $\pi/2$. As a result, the energy spectrum of the particle is always bounded from below *and* from above, and it ranges from m to $\pi/2$ for particles and from $-m$ to $-\pi/2$ for antiparticles.

A peculiar situation arises when $m = \pi/2$. The energy ρ is then independent of the velocity of the particle. In this case, the deficit angle is always equal to π , so that, referring again to figure 2, exactly half of the Minkowski space is cut out. Clearly, every observer will then agree with this fact. It is also the limit that cannot be transgressed by accelerating a particle with a mass different from $\pi/2$. Hence, if the mass is smaller than $\pi/2$ (or bigger, for an antiparticle), then the deficit angle is always smaller than $\pi/2$, and vice versa.

Lightlike particles

Finally, let us briefly come back to the two types of lightlike particles, with $m = 0$ and $m = \pi$, and let us, for simplicity, consider the particle states only. The antiparticles behave similarly, if one replaces $m = 0$ with $m = \pi$ and vice versa. For a particle with $m = 0$, the mass shell condition implies that $0 < \rho < \pi/2$, and thus the deficit angle lies in the range $0 < \alpha < \pi$. This is the situation shown in figure 2, if we take p to be a lightlike vector. There is nothing special about this spacetime, except that it is not possible to transform to a rest frame of the particle. The observer simply sees a particle moving with the speed of light.

Something rather different happens for $m = \pi$. Then the deficit angle lies in the range $\pi < \alpha < 2\pi$, and the two planes in figure 2 point to the right, into the direction in which the particle is moving with the speed of light. The crucial difference to the former situation is that it is now, for a test particle moving on a straight timelike line, not possible to “escape” the planes θ_\pm . After some finite time, it will be hit by one of

them, say θ_- , and reappear at θ_+ . Then, again after some finite time, the same happens again, and so on. One can show that the intervals between these events become smaller and smaller, and that the test particle will finally collide with the lightlike particle.

Expressed differently, a particle with $m = \pi$ is a rather strong gravitational attractor, although there is no local gravitational force. But nevertheless, everything moving on timelike lines will hit the attractor after some finite time. It is almost something like a “big crunch” singularity of the universe, where every timelike and every lightlike geodesic ends on, except that lightlike geodesics can still escape the singularity. For a different way to show that this is a typical behaviour of a lightlike conical singularity, we refer to section 3.4 in [14]. There, it is also shown that for spacelike holonomies \mathbf{u} , the singularity essentially becomes a big crunch (or big bang) singularity. This is another way to see that $m = \pi$ is an upper bound for the mass. The conical singularity on the world line of the particle is then already something in between a matter source and a big crunch. It is the strongest possible gravitational attractor which is located on a causal world line.

3 Hamiltonian formulation

In this section, we want to derive the classical Hamiltonian formalism for the point particle, coupled to Einstein gravity as a field theory. What we need for this purpose is a proper definition of the configuration space of the system, and an action principle. The action should be given by the Einstein Hilbert action for the gravitational fields, plus an appropriate term for the particle. We shall also need some fall off conditions and boundary terms at infinity, to make the action finite and to get the correct equations of motion. Furthermore, it turns out that we have to add some additional degrees of freedom to the model, and these can most naturally be interpreted as an observer at infinity.

All this is most conveniently carried out if we switch to the ADM formulation from the very beginning. We split our spacetime manifold into $\mathcal{M} = \mathbb{R} \times \mathcal{N}$, where $\mathcal{N} = \mathbb{R}_+ \times S^1$ is the space manifold. It has a circular boundary, which represents the position of the particle at a given moment of time. On \mathcal{N} , we introduce formal tangent indices i, j, \dots , taking the values r and φ , a Levi Civita tensor ε^{ij} with $\varepsilon^{r\varphi} = 1$, and we denote derivatives with respect to t by a dot. The one-forms e_μ and ω_μ on \mathcal{M} split into one-forms e_i and ω_i on \mathcal{N} , and scalars e_t and ω_t . The Einstein Hilbert action becomes the time integral of the Lagrangian

$$L = \int_{\mathcal{N}} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\dot{\omega}_i e_j + \omega_t D_i e_j + \frac{1}{2} e_t F_{ij}), \quad (3.1)$$

and the vacuum Einstein equations split into *constraints*

$$\varepsilon^{ij} F_{ij} = 0, \quad \varepsilon^{ij} D_i e_j = 0, \quad (3.2)$$

and *evolution equations*

$$\dot{\omega}_i = D_i \omega_t, \quad \dot{e}_i = D_i e_t + [e_i, \omega_t]. \quad (3.3)$$

In (3.1), we already performed an integration by parts, such that the derivatives do not act on the time components of the fields. As we have to add boundary terms anyway, we may equally well use this as an ansatz for the total Lagrangian. It has the advantage that we can immediately read off the constraints by varying L with respect to the “multipliers” e_t and ω_t . To derive the evolution equations, however, we have to take into account a boundary term. Varying the Lagrangian with respect to the “dynamical” fields e_i and ω_i yields

$$\begin{aligned} \delta L = & \int_{\mathcal{N}} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}((\dot{\omega}_i - D_i \omega_t) \delta e_j + (\dot{e}_i - D_i e_t - [e_i, \omega_t]) \delta \omega_j) \\ & + \int_{\mathcal{N}} d^2x \varepsilon^{ij} \frac{1}{2} \partial_i \text{Tr}(\omega_t \delta e_j + e_t \delta \omega_j). \end{aligned} \quad (3.4)$$

To be precise, there is also a total time derivative that has been neglected, but this is of course allowed, because we are only interested in the Euler Lagrange equations. From the first line in this expression, we deduce the evolution equations. The second line can be converted into a boundary integral. There are two boundaries of \mathcal{N} that contribute, namely the “particle boundary” at $r = 0$, and the “boundary at infinity”, where $r \rightarrow \infty$. For a generic one-form ξ_i on \mathcal{N} , we have, according to Stokes’ theorem,

$$\int_{\mathcal{N}} d^2x \varepsilon^{ij} \partial_i \xi_j = \int_{r \rightarrow \infty} d\varphi \xi_\varphi - \int_{r=0} d\varphi \bar{\xi}_\varphi. \quad (3.5)$$

Let us, for the moment, ignore the boundary at infinity, which will be considered in detail later on. The contribution to δL from the particle boundary is then given by

$$- \int_{r=0} d\varphi \frac{1}{2} \text{Tr}(\bar{\omega}_t \delta \bar{e}_\varphi + \bar{e}_t \delta \bar{\omega}_\varphi). \quad (3.6)$$

As this has to vanish for a generic variation of the fields, we get additional equations of motion on the boundary, namely $\bar{e}_t = 0$ and $\bar{\omega}_t = 0$. Clearly, this is not what we want. Instead, we want the point particle condition and the mass shell condition to be imposed.

The particle action

As an ansatz, let us introduce Lagrange multipliers ζ and $\bar{\chi}$, and add a constraint term to the Lagrangian

$$L_{\text{prt}} = -\zeta(\frac{1}{2} \text{Tr}(\bar{\mathbf{u}}) - \cos m) - \int_{r=0} d\varphi \frac{1}{2} \text{Tr}(\bar{\chi} \bar{e}_\varphi). \quad (3.7)$$

Note that, to impose the point particle condition, we need an $\mathfrak{sl}(2)$ valued field $\bar{\chi}$ on the boundary, but for the mass shell constraint it is sufficient to introduce a single real variable ζ . This is because the trace of the holonomy of the particle is, by definition, independent of φ and thus, at a given time t , the mass shell condition is only a single real equation.

If we now vary L_{prt} with respect to the multipliers, we get the correct constraints. But we also get a contribution from L_{prt} to the variation with respect to the gravitational fields. To compute this, we need the derivative of the holonomy (2.13) with respect to the spin connection, which is given by

$$\frac{1}{2} \delta \text{Tr}(\bar{\mathbf{u}}) = - \int d\varphi \frac{1}{2} \text{Tr}(\bar{\mathbf{u}} \delta \bar{\boldsymbol{\omega}}_\varphi) = - \int d\varphi \frac{1}{2} \text{Tr}(\bar{\mathbf{p}} \delta \bar{\boldsymbol{\omega}}_\varphi). \quad (3.8)$$

The last equality holds because $\bar{\boldsymbol{\omega}}_\varphi$ is traceless, so that only the projection $\bar{\mathbf{p}}$ of the holonomy $\bar{\mathbf{u}}$ contributes to the trace. Using this, we find

$$\delta L_{\text{prt}} = \int_{r=0} d\varphi \frac{1}{2} \text{Tr}(\zeta \bar{\mathbf{p}} \delta \bar{\boldsymbol{\omega}}_\varphi - \bar{\chi} \delta \mathbf{e}_\varphi), \quad (3.9)$$

and adding this to (3.6), the extra field equations on the boundary become

$$\bar{\boldsymbol{\omega}}_t = -\bar{\chi}, \quad \bar{\mathbf{e}}_t = \zeta \bar{\mathbf{p}}. \quad (3.10)$$

Now, the first equation only fixes the values of the multiplier field $\bar{\chi}$. We can get rid of this, if we identify it from the very beginning with $\bar{\boldsymbol{\omega}}_t$. Hence, we redefine the particle Lagrangian to be

$$L_{\text{prt}} = -\zeta \left(\frac{1}{2} \text{Tr}(\bar{\mathbf{u}}) - \cos m \right) + \int_{r=0} d\varphi \frac{1}{2} \text{Tr}(\bar{\boldsymbol{\omega}}_t \bar{\mathbf{e}}_\varphi). \quad (3.11)$$

We are then left with the second equation in (3.10). Except for the fact that it fixes the value of the multiplier ζ , this turns out to be redundant. It is a “secondary constraint”, which does not impose any further restriction on the fields. One can show explicitly that it follows from the point particle condition and its time derivative, but we can also use the results of the previous section. There we found, provided that the Einstein equations are satisfied, that the momentum of the particle is given by (2.15), and that the time component of the dreibein is given by a similar expression (2.11),

$$\bar{\mathbf{p}} = \bar{\mathbf{g}}^{-1} \mathbf{p} \bar{\mathbf{g}}, \quad \bar{\mathbf{e}}_t = \partial_t \tau \bar{\mathbf{g}}^{-1} \mathbf{p} \bar{\mathbf{g}}. \quad (3.12)$$

Obviously, this implies that the time component of the dreibein on the boundary is proportional to the representation of the momentum vector in the local frames. Or, that the velocity of the particle is proportional to its momentum. The only extra information contained in (3.10) is that the physical time elapsing on the world line of the particle is determined by the value of the multiplier ζ of the mass shell constraint.

The fall off condition

Let us now return to the “boundary at infinity”. The problem is somewhat more complicated there, because, as it stands, the Lagrangian will in general not converge at $r \rightarrow \infty$. Let us write it explicitly as a limit, ignoring the particle term for a moment,

$$L = \lim_{R \rightarrow \infty} L_{\text{int}}, \quad L_{\text{int}} = \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\dot{\boldsymbol{\omega}}_i \mathbf{e}_j + \boldsymbol{\omega}_t D_i \mathbf{e}_j + \frac{1}{2} \mathbf{e}_t \mathbf{F}_{ij}). \quad (3.13)$$

We can now look for a minimal condition for this to converge, demanding some additional properties of this fall off condition. Thereby we shall, somewhat implicitly, use that we are dealing with a topological field theory. As there are no local physical degrees of freedom carried by the gravitational field, it will be possible to localize the total energy, or the total action of a physical field configuration, that is, a solution to the field equations, within a compact region of the space manifold \mathcal{N} . As a consequence, the total action of any field configuration satisfying the Einstein equations is always finite. We do not need to impose any *further* restrictions. All we need to do, in order to get a well defined action principle, is to restrict the *off shell* field configurations to those that still have a finite action.

If we also demand that the condition has the usual features, namely, that it only restricts the behaviour of the fields at infinity, and that it does not involve time derivatives, then it is, quite surprisingly, already fixed. It must be a “subset” (in the logical sense) of the field equations, as otherwise it would restrict their solutions. Actually, it must be a subset of the constraints, because the other field equations contain time derivatives. But then it cannot impose any restriction on the multipliers e_t and ω_t , because these are not involved in the constraints.

Now, consider the last two terms in the integrand above, which are proportional to the multipliers. If we want them to converge for *any* choice of e_t and ω_t , then the constraints themselves must have a compact support. Hence, the fall off condition must read

$$\exists R \quad \text{such that, for } r \geq R, \quad \varepsilon^{ij} \mathbf{F}_{ij} = 0, \quad \varepsilon^{ij} D_i \mathbf{e}_j = 0. \quad (3.14)$$

We can write this in a slightly different way. Similar to the full Einstein equations, the general solution to the constraints can be given in terms of an embedding (g, f) . It is formally the same as (1.11), except that now we have to replace the spacetime indices μ by space indices i . And, as in the previous section, we have to cut the region $r \geq R$, which has the shape of a ring, along some radial line, say, again at $\varphi = \pm\pi$. As a result, the fall off condition becomes

$$\exists R \quad \text{such that, for } r \geq R, \quad \omega_i = g^{-1} \partial_i g, \quad e_i = g^{-1} \partial_i f g. \quad (3.15)$$

Along the cut at $\varphi = \pm\pi$, the fields g and f are again denoted by g_{\pm} and f_{\pm} , and they have to satisfy a boundary condition like (2.6),

$$g_+ = u^{-1} g_-, \quad f_+ = u^{-1} (f_- - v) u. \quad (3.16)$$

Note, however, that the situation is somewhat different from the one in the previous section. It is now only the space manifold \mathcal{N} , and in fact only a “neighbourhood of infinity”, which is embedded into Minkowski space. The fall off condition only requires that such an embedding exists at each moment of time, but it does not say anything about how the embedding at one time is related to that at another time. In particular, the “constants” $u \in \text{SL}(2)$ and $v \in \mathfrak{sl}(2)$ can be arbitrary functions of time.

There is also a small but important difference between the two formulations of the fall off condition, (3.14) and (3.15). Both can be thought of as a kind of “asymptotical flatness” condition. They state that it is possible to embed a neighbourhood of spatial

infinity into flat Minkowski space. But in the second form, such an embedding is given *explicitly*. Now, remember that the embedding (\mathbf{g}, \mathbf{f}) was determined by the dreibein and the spin connection up to a rigid Poincaré transformation only. Therefore, the fields \mathbf{g} and \mathbf{f} contain more information than the dreibein and the spin connection.

If we use the second form (3.15) as the fall off condition, and regard \mathbf{g} and \mathbf{f} as additional variables, then we add six new degrees of freedom to the system, corresponding to the six dimensions of the Poincaré group. At this stage, these are pure gauge degrees of freedom, because the action does not depend on them. However, we shall see in a moment that, in order to get a finite action, we are forced to add boundary terms to the action. They will explicitly depend on \mathbf{g} and \mathbf{f} . As a consequence, the six new degrees of freedom associated with the embedding will become physical. The action will no longer be gauge invariant under Poincaré transformations of the embedding Minkowski space.

There is also a natural physical interpretation of these extra degrees of freedom, which is typical for all kinds of asymptotical flatness conditions. We can think of the background frame at infinity as the rest frame of some observer. What we are effectively doing if we consider the embedding of spatial infinity as an additional degree of freedom of our model, is to add an external observer to the universe. A Poincaré transformation acting on \mathbf{g} and \mathbf{f} then corresponds to a combination of rotations, boosts, and translations of the observer's reference frame. So far, it is however not clear whether this interpretation makes sense. What we have to show is that the observer "couples" correctly to the rest of the universe. In particular, the reference frame should stay fixed under the time evolution, but so far we do not have any equations of motion for the embedding.

The boundary action at infinity

To render the kinetic term in the Lagrangian (3.13) finite, we cannot impose any further restriction on the fields at infinity, because we would then restrict the solutions to the field equations. The only way out is to regularize the action by adding a compensating boundary term. It is not obvious that such a term exists, but we shall see in a moment that it does. Let us compute the derivative of L_{int} with respect to R , assuming that R is large enough, such that the fall off condition holds for this R . Of course, it then also holds for every larger R , and we can neglect the constraint terms,

$$\frac{\partial L_{\text{int}}}{\partial R} = \frac{\partial}{\partial R} \left(\int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\dot{\omega}_i e_j) \right) = \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\dot{\omega}_r e_\varphi - \dot{\omega}_\varphi e_r). \quad (3.17)$$

Inserting (3.15), and using the general formula $\partial(\mathbf{g}^{-1} \delta \mathbf{g}) = \mathbf{g}^{-1} \delta(\partial \mathbf{g} \mathbf{g}^{-1}) \mathbf{g}$, which holds for any two commuting derivations ∂ and δ , this becomes

$$\frac{\partial L_{\text{int}}}{\partial R} = \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\partial_r(\dot{\mathbf{g}} \mathbf{g}^{-1}) \partial_\varphi \mathbf{f} - \partial_\varphi(\dot{\mathbf{g}} \mathbf{g}^{-1}) \partial_r \mathbf{f}). \quad (3.18)$$

Here and in the following the integral over φ always runs from $-\pi$ to π . If we now integrate the derivative ∂_φ in the last term by parts, we get

$$\frac{\partial L_{\text{int}}}{\partial R} = \int_{r=R} d\varphi \frac{1}{2} \partial_r \text{Tr}(\dot{\mathbf{g}} \mathbf{g}^{-1} \partial_\varphi \mathbf{f}) + \frac{1}{2} \text{Tr}(\dot{\mathbf{g}}_- \mathbf{g}_-^{-1} \partial_r \mathbf{f}_- - \dot{\mathbf{g}}_+ \mathbf{g}_+^{-1} \partial_r \mathbf{f}_+). \quad (3.19)$$

Using the boundary condition (3.16), this can be simplified to

$$\frac{\partial L_{\text{int}}}{\partial R} = \frac{\partial}{\partial R} \left(\int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\dot{\mathbf{g}} \mathbf{g}^{-1} \partial_\varphi \mathbf{f}) + \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \dot{\mathbf{u}} \mathbf{f}_+) \right). \quad (3.20)$$

From this we infer that the action becomes finite if we add the following boundary term,

$$L_{\text{bnd}} = - \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\dot{\mathbf{g}} \mathbf{g}^{-1} \partial_\varphi \mathbf{f}) - \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \dot{\mathbf{u}} \mathbf{f}_+). \quad (3.21)$$

Note that this is a kind of ‘‘Wess Zumino Witten’’ term that typically appears as a boundary action of a topological field theory [19]. Taking everything together, the total Lagrangian becomes

$$L = \lim_{R \rightarrow \infty} (L_{\text{prt}} + L_{\text{int}} + L_{\text{bnd}}), \quad (3.22)$$

and it is now manifestly finite, because the fall off condition implies that, from some minimal R on, the expression between the parenthesis no longer depends on R . The convergence is therefore quite strong, and we can evaluate the Lagrangian by taking R large enough. We can however not drop the limit, because the minimal value for R depends on the field configuration.

The phase space

We now have a well defined Lagrangian for our model, but so far we haven’t shown that it provides the correct equations of motion at spatial infinity. This is somewhat involved, mainly due to the relations between the gravitational fields and the embedding at spatial infinity. The fall off condition restricts the possible variations of the fields in a non-trivial way, and these restrictions have to be taken into account when we derive the equations of motion. The actual calculation is most convenient carried out at the phase space level. We can then also use some of the results later on for the phase space reduction.

From the first order structure of the Lagrangian, it is more or less obvious how the basic fields split into phase space variables and Lagrange multipliers. The phase space is spanned by the spatial components of the gravitational fields, e_i and ω_i , and the embedding fields at spatial infinity, \mathbf{g} and \mathbf{f} . Let us denote these fields collectively by Φ , and call the set of all such field configurations, subject to the ADM condition and the fall off condition, the *extended phase space* $\tilde{\mathcal{P}}$. From this, we shall later derive the *reduced phase space* \mathcal{P} as the quotient space of $\tilde{\mathcal{P}}$ modulo gauge transformations. The total Lagrangian can then be split into a kinetic term,

$$L_{\text{kin}} = \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\dot{\omega}_i e_j) - \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\dot{\mathbf{g}} \mathbf{g}^{-1} \partial_\varphi \mathbf{f}) - \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \dot{\mathbf{u}} \mathbf{f}_+), \quad (3.23)$$

and the Hamiltonian

$$\begin{aligned} H = \zeta \left(\frac{1}{2} \text{Tr}(\bar{\mathbf{u}}) - \cos m \right) &- \int_{r=0} d\varphi \frac{1}{2} \text{Tr}(\bar{\omega}_t \bar{e}_\varphi) \\ &- \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\omega_t D_i e_j + \frac{1}{2} e_t F_{ij}). \end{aligned} \quad (3.24)$$

Here, we did not write out the limit $R \rightarrow \infty$ explicitly, but it has to be understood in all what follows. The kinetic Lagrangian is thus a function of the phase space variables Φ and their velocities $\dot{\Phi}$. As it is linear in the velocities, we can directly read off the symplectic potential on $\tilde{\mathcal{P}}$. It is given by

$$\Theta = \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\mathbf{d}\omega_i \mathbf{e}_j) - \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\mathbf{d}\mathbf{g} \mathbf{g}^{-1} \partial_\varphi \mathbf{f}) - \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \mathbf{d}\mathbf{u} \mathbf{f}_+), \quad (3.25)$$

where \mathbf{d} denotes the exterior derivative on $\tilde{\mathcal{P}}$. Acting with another exterior derivative on this, we find the symplectic form

$$\begin{aligned} \Omega = \mathbf{d}\Theta = & - \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\mathbf{d}\omega_i \mathbf{d}\mathbf{e}_j) \\ & + \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\mathbf{d}\mathbf{g} \mathbf{g}^{-1} \partial_\varphi \mathbf{d}\mathbf{f} - \mathbf{d}\mathbf{g} \mathbf{g}^{-1} \mathbf{d}\mathbf{g} \mathbf{g}^{-1} \partial_\varphi \mathbf{f}) \\ & + \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \mathbf{d}\mathbf{u} \mathbf{u}^{-1} \mathbf{d}\mathbf{u} \mathbf{f}_+ + \mathbf{u}^{-1} \mathbf{d}\mathbf{u} \mathbf{d}\mathbf{f}_+). \end{aligned} \quad (3.26)$$

This looks somewhat complicated, but, as we shall see in a moment, it is a non-degenerate two-form on $\tilde{\mathcal{P}}$. We could therefore, in principle, derive the Poisson bracket from it. But this turns out to be rather difficult. For example, the brackets of the fields \mathbf{g} and \mathbf{f} will be non-local (see, e.g., the appendix in [14], where a very similar action is considered). It is much more convenient to stick to the symplectic structure as the basic object on the phase space, and give an indirect proof of the non-degeneracy. One way to do this is to show that the Hamiltonian evolution equations,

$$\Omega \cdot \dot{\Phi} = \mathbf{d}H, \quad (3.27)$$

can be solved uniquely for the velocities of the fields $\Phi \in \tilde{\mathcal{P}}$. The left hand side of this equation is obtained by contracting the symplectic form with the velocities,

$$\begin{aligned} \Omega \cdot \dot{\Phi} = & \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\dot{\omega}_i \mathbf{d}\mathbf{e}_j + \dot{\mathbf{e}}_i \mathbf{d}\omega_j) \\ & + \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(\dot{\omega}_\varphi \mathbf{g}^{-1} \mathbf{d}\mathbf{f} \mathbf{g} + \dot{\mathbf{e}}_\varphi \mathbf{g}^{-1} \mathbf{d}\mathbf{g}) \\ & + \frac{1}{2} \text{Tr}(\dot{\mathbf{g}}_- \mathbf{g}_-^{-1} \mathbf{d}\mathbf{v} + (\dot{\mathbf{f}}_+ + [\mathbf{f}_+, \dot{\mathbf{g}}_+ \mathbf{g}_+^{-1}]) \mathbf{u}^{-1} \mathbf{d}\mathbf{u}). \end{aligned} \quad (3.28)$$

Here, we integrated several times by parts, made use of the boundary condition (3.16) along the cut at $\varphi = \pm\pi$, and we used that, for sufficiently large r , the derivatives of ω_i and \mathbf{e}_i are given in terms of those of \mathbf{g} and \mathbf{f} by

$$\begin{aligned} \mathbf{d}\omega_i &= \mathbf{d}(\mathbf{g}^{-1} \partial_i \mathbf{g}) = \mathbf{g}^{-1} \partial_i (\mathbf{d}\mathbf{g} \mathbf{g}^{-1}) \mathbf{g}, \\ \mathbf{d}\mathbf{e}_i &= \mathbf{d}(\mathbf{g}^{-1} \partial_i \mathbf{f} \mathbf{g}) = \mathbf{g}^{-1} \partial_i \mathbf{d}\mathbf{f} \mathbf{g} + [\mathbf{g}^{-1} \partial_i \mathbf{f} \mathbf{g}, \mathbf{g}^{-1} \mathbf{d}\mathbf{g}]. \end{aligned} \quad (3.29)$$

Similarly, the exterior derivative of the Hamiltonian is found to be

$$\mathbf{d}H = \int_{r=0} d\varphi \frac{1}{2} \text{Tr}((\bar{\mathbf{e}}_t - \zeta \bar{\mathbf{p}}) \mathbf{d}\omega_\varphi)$$

$$\begin{aligned}
& + \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(D_i \omega_t d\mathbf{e}_j + (D_i \mathbf{e}_t + [\mathbf{e}_i, \omega_t]) d\omega_j) \\
& + \int_{r=R} d\varphi \frac{1}{2} \text{Tr}(D_\varphi \omega_t \mathbf{g}^{-1} d\mathbf{f} \mathbf{g} + (D_\varphi \mathbf{e}_t + [\mathbf{e}_s, \omega_t]) \mathbf{g}^{-1} d\mathbf{g}) \\
& + \frac{1}{2} \text{Tr}(\mathbf{g}_- \omega_t \mathbf{g}_-^{-1} d\mathbf{v} + (\mathbf{g}_+ \mathbf{e}_t \mathbf{g}_+^{-1} + [\mathbf{f}_+, \mathbf{g}_+ \omega_t \mathbf{g}_+^{-1}]) \mathbf{u}^{-1} d\mathbf{u}). \quad (3.30)
\end{aligned}$$

Now, we can read off the evolution equations, and the secondary constraints on the particle boundary. The latter follows from the first line in (3.30), which has no counterpart in (3.28). It is exactly the second equation in (3.10). To recover the evolution equations for ω_i and e_i , we have to compare the integrals over the range $r \leq R$ in (3.28) and (3.30). Obviously, they are still the same as (3.3), and they must be satisfied *everywhere* on \mathcal{N} , because the two integrals must be equal in the limit $R \rightarrow \infty$.

To find the remaining equations of motion, let us assume that those already obtained are satisfied. We are then left with the last lines in (3.28) and (3.30), containing terms proportional to the exterior derivatives of \mathbf{u} and \mathbf{v} . They also contain the time derivatives of the fields \mathbf{g} and \mathbf{f} , so we expect them to provide an evolution equation for the embedding. Now, the time evolution of the embedding is already almost fixed by that of the dreibein and the spin connection. The only remaining freedom is to perform a rigid Poincaré transformation. Explicitly, we find that

$$\dot{\mathbf{g}} = \mathbf{g} \omega_t - \mathbf{c} \mathbf{g}, \quad \dot{\mathbf{f}} = \mathbf{g} \mathbf{e}_t \mathbf{g}^{-1} - [\mathbf{c}, \mathbf{f}] - \mathbf{d}, \quad (3.31)$$

where the parameters $\mathbf{c}, \mathbf{d} \in \mathfrak{sl}(2)$ are spatially constant, but may be chosen freely at each moment of time. This freedom corresponds to an infinitesimal generator of a Poincaré transformation (1.15). Without the boundary terms in the action, \mathbf{c} and \mathbf{d} would provide six free parameters, and the embedding would be a gauge degree of freedom. But now we get six extra equations of motion as well, which fix the constants \mathbf{c} and \mathbf{d} . Inserting (3.31), the difference of the last lines in (3.28) and (3.30) becomes a rather simply expression,

$$\frac{1}{2} \text{Tr}(\mathbf{c} d\mathbf{v} + \mathbf{d} \mathbf{u}^{-1} d\mathbf{u}). \quad (3.32)$$

This vanishes if and only if $\mathbf{c} = 0$ and $\mathbf{d} = 0$, because \mathbf{u} and \mathbf{v} are *independent functions* on the phase space $\tilde{\mathcal{P}}$. They are implicitly defined by the relations (3.16), and it is easy to check that their values are not restricted. The time evolution of the embedding of spatial infinity is therefore fixed, and we have

$$\mathbf{g}^{-1} \dot{\mathbf{g}} = \omega_t, \quad \mathbf{g}^{-1} \dot{\mathbf{f}} \mathbf{g} = \mathbf{e}_t. \quad (3.33)$$

Together with the fall off condition (3.15), we recover the three dimensional equations (1.11). As a result, the frame at infinity behaves such that it provides an embedding of *spacetime* into Minkowski space, although it has been defined as an embedding of *space* only. We can also say that the evolution equations are such that the observer at infinity is correctly “coupled” to the universe. His reference frame is not shifted or rotated under the time evolution.

Moreover, we were able to solve the Hamiltonian equations of motion uniquely for the velocities of the phase space variables $\dot{\Phi}$, for any initial field configuration Φ . This

implies that the symplectic form (3.26) is non-degenerate, as otherwise the Hamiltonian equations (3.27) could not be solved for $\dot{\Phi}$. Without explicitly deriving the inverse of Ω , that is, the Poisson bracket, we know that we identified the phase space structure of the system correctly.

Gauge symmetries

Finally, let us very briefly discuss the gauge symmetries of our model, which we want to divide out in the next section. The basic gauge symmetries of Einstein gravity in the dreibein formulation are the local Poincaré transformations,

$$\delta\omega_\mu = D_\mu\lambda, \quad \delta e_\mu = D_\mu\xi + [e_\mu, \lambda]. \quad (3.34)$$

Here, $\lambda \in \mathfrak{sl}(2)$ is the generator of a local Lorentz rotation, and $\xi \in \mathfrak{sl}(2)$ that of a local translation. For the fall off condition to be preserved, the embedding fields have to transform according to

$$\delta g = g\lambda, \quad \delta f = g\xi g^{-1}. \quad (3.35)$$

The generator of a spacetime diffeomorphism is obtained, at least on shell, by fixing a vector field ξ^μ and choosing the parameters to be $\lambda = \xi^\mu\omega_\mu$ and $\xi = \xi^\mu e_\mu$. Then we have

$$\begin{aligned} \delta e_\mu &= \xi^\nu \partial_\nu e_\mu + \partial_\mu \xi^\nu e_\nu + 2\xi^\nu D_{[\mu} e_{\nu]}, \\ \delta\omega_\mu &= \xi^\nu \partial_\nu \omega_\mu + \partial_\mu \xi^\nu \omega_\nu + \xi^\nu F_{\mu\nu}. \end{aligned} \quad (3.36)$$

Up to terms proportional to the field equations, these are the Lie derivatives of the one-forms. Similarly, one finds that the same holds for the scalars g and f . It is also straightforward to check that all these transformations are symmetries of the action, provided that, on the particle boundary, the parameter ξ obeys

$$\bar{\xi} = \eta \bar{p}, \quad (3.37)$$

for some real number η , and that one includes a transformation of the multiplier of the mass shell constraint,

$$\delta\zeta = \dot{\eta}. \quad (3.38)$$

Of course, this restriction arises in the same way as the secondary constraint for the time component of the dreibein on the particle boundary (3.10). It is a result of the restricted diffeomorphisms invariance on the boundary. If we write (3.37) as a restriction on the generating vector field ξ^μ of a diffeomorphism, then the condition becomes $\bar{\xi}^r = 0$ and $\partial_\varphi \bar{\xi}^t = 0$. The first condition is quite reasonable. It states that the vector field is tangent to the boundary, which means that the diffeomorphism maps the boundary onto itself. The second condition implies that the diffeomorphism does not mix the circles of constant t on the boundary.

In other words, we do not have the full diffeomorphism group on the boundary as a gauge group. This is because the circles of constant t play a distinguished role. They

represent a *single physical point* on the world line of the particle. It is the point particle equation $\bar{e}_\varphi = 0$, which explicitly refers to these circles, and brakes the diffeomorphism invariance. Off the boundary however the diffeomorphism group is unbroken, and there is a one-to-one relation between local translations and diffeomorphisms, because the dreibein, which relates the generating vector fields ξ^μ and ξ , is invertible. We should also note that the fall off condition is diffeomorphism invariant, although it refers to the special coordinate r . It is equivalent to the manifestly invariant statement that there exists an embedding outside some compact subset of \mathcal{N} .

At the phase space level, the local Poincaré transformations behave as it is expected. The transformation of the phase space variables e_i , ω_i , g and f are determined by the parameters λ and ξ alone, whereas those of the multipliers e_t , ω_t and ζ depend on the time derivatives of the parameters as well. We can also show that the gauge transformations $\delta\Phi$ on the phase space $\tilde{\mathcal{P}}$ are generated by the constraints. We have $\delta\Phi = \{C, \Phi\}$, or equivalently $\Omega \cdot \delta\Phi = dC$, if we set

$$C = \eta \left(\frac{1}{2} \text{Tr}(\bar{u}) - \cos m \right) - \int_{r=0} d\varphi \frac{1}{2} \text{Tr}(\bar{\lambda} \bar{e}_\varphi) - \int_{r \leq R} d^2x \varepsilon^{ij} \frac{1}{2} \text{Tr}(\lambda D_i e_j + \frac{1}{2} \xi F_{ij}). \quad (3.39)$$

This is, like the Hamiltonian, a linear combination of the constraints. But it is not the most general linear combination. The parameter of local translations on the boundary is restricted by (3.37). It is however the most general linear combination that is *first class* and thus a generator of a gauge symmetry.

That some of the constraints are second class can also be inferred from the fact that there are secondary constraints which restrict the associated multipliers, and therefore the gauge symmetries. An explicit derivation of the Poisson bracket on the extended phase space $\tilde{\mathcal{P}}$ would therefore not be of much use. For the derivation of the reduced phase space, we had to go over to the Dirac bracket anyway, and this would make things even more complicated. If we however stick to the symplectic formalism, the corresponding structure on the reduced phase space can be found quite easily, by pulling back the symplectic form Ω onto the constraint surface, and then dividing out the gauge degrees of freedom, which happens almost automatically.

Finally, let us briefly come back to the gauge transformations (1.7) defined in the beginning as the most general transformations of the local frames that leave the metric and the time orientation of spacetime invariant,

$$e_\mu \mapsto k^{-1} e_\mu k, \quad \omega_\mu \mapsto k^{-1} (\partial_\mu + \omega_\mu) k. \quad (3.40)$$

This is the finite version of a local Lorentz transformation. There is however a crucial difference between the infinitesimal transformations above and these finite ones. Some of the transformation given by (3.40) are *large* gauge transformations, which means that they are not generated by (3.34).

As an example, consider the maps $k = e^{-z\varphi\gamma_0}$, with $z \in \mathbb{Z}$, as parameters of local Lorentz transformations. They are well defined, because $e^{2\pi\gamma_0} = 1$, and one can easily show that every map $k : \mathcal{N} \rightarrow \text{SL}(2)$ can be smoothly deformed into one of these, but it is not possible to deform them into each other. So, the set of all maps k ,

and therefore the gauge group, splits into a series of disconnected components, labeled by the “winding number” z . Only the transformations with $z = 0$ are generated by the infinitesimal transformations (3.34).

To see what a large gauge transformations does, let us choose $\mathbf{k} = e^{-z\varphi\gamma_0}$ and act with this on the fields in the neighbourhood of the particle, as given by (2.2) and (2.3). The result is that the dreibein is replaced by

$$\begin{aligned} e_t &= \gamma_0, & e_r &= \cos((2z+1)\varphi) \gamma_1 + \sin((2z+1)\varphi) \gamma_2, \\ e_\varphi &= \left(1 - \frac{\pi}{m}\right) r \left(\cos((2z+1)\varphi) \gamma_2 - \sin((2z+1)\varphi) \gamma_1 \right), \end{aligned} \quad (3.41)$$

and the spin connection becomes

$$\omega_t = 0, \quad \omega_r = 0, \quad \omega_\varphi = -\left(\frac{m}{2\pi} + z\right) \gamma_0. \quad (3.42)$$

So, the new dreibein “winds” differently around the boundary at $r = 0$, and the value of the angular component of the spin connection has changed as well. But, for example, the holonomy of the particle is still the same,

$$\bar{\mathbf{u}} = e^{(m+2\pi z)\gamma_0} = e^{m\gamma_0}, \quad (3.43)$$

and the metric is of course also unchanged. There is no way to distinguish the field configurations for different winding numbers z physically, if we are only allowed to refer to the metric, and to use vector or spinor test particles to measure holonomies. This is the reason why we have to consider the large Lorentz transformations as gauge transformations. What seems a bit strange, however, is that the winding number of the dreibein in (3.41) is always odd. The following dreibein represents the same metric,

$$\begin{aligned} e_t &= \gamma_0, & e_r &= \cos(2z\varphi) \gamma_1 + \sin(2z\varphi) \gamma_2, \\ e_\varphi &= \left(1 - \frac{\pi}{m}\right) r \left(\cos(2z\varphi) \gamma_2 - \sin(2z\varphi) \gamma_1 \right), \end{aligned} \quad (3.44)$$

but it is not possible to transform it into (3.41). What is the difference between these field configurations? To find the answer, we have to evaluate the spin connection,

$$\omega_t = 0, \quad \omega_r = 0, \quad \omega_\varphi = -\left(\frac{m-\pi}{2\pi} + z\right) \gamma_0, \quad (3.45)$$

and compute the holonomy

$$\bar{\mathbf{u}} = e^{(m-\pi+2\pi z)\gamma_0} = e^{-(\pi-m)\gamma_0} = -e^{m\gamma_0}. \quad (3.46)$$

This is not the same as (3.43). The two particles can be distinguished physically, by transporting a spinor around them. The resulting spinors will differ by an overall sign. The second particle is, using the terminology of the previous section, an *antiparticle* with mass $\pi - m$. We have seen that this has the same deficit angle as a particle with mass m , so that the metric in the neighbourhood is the same.

From this we conclude, and this will be rather important later on, that the whole physical information about the particle is encoded in the *spinor* representation of the

holonomy $\bar{u} \in \text{SL}(2)$. It is not sufficient to consider the vector representation, because then we cannot distinguish between particles and antiparticles. But on the other hand, it is also not appropriate to consider its representation in any “higher” covering of $\text{SL}(2)$ as a physical quantity. This would no longer be invariant under general local Lorentz transformations, because the last equality in (3.43) no longer holds. This is why $\text{SL}(2)$ will become to the “momentum space” of the particle in the next section, and not a covering thereof.

4 Phase space reduction

When quantizing a gauge theory, there are two roads between which one has to choose. Either one can quantize the full theory, including gauge degrees of freedom, and use operator constraint equations on the Hilbert space, or one can divide out the gauge degrees of freedom at the classical level, and then quantize. In the case of a topological field theory, the number of physical degrees of freedom that remain after such a phase space reduction is finite, and therefore it is more convenient to first reduce and then quantize. This is the road we will follow in this article. We shall however, in a first step, not divide out all gauge symmetries. Instead, we only consider those that are generated by the “kinematical” constraints,

$$\varepsilon^{ij} \mathbf{F}_{ij} = 0, \quad \varepsilon^{ij} D_i \mathbf{e}_j = 0, \quad \bar{e}_\varphi = 0. \quad (4.1)$$

If we do so, then there is only one constraint left, the mass shell constraint, and this will become a “dynamical” constraint, in the sense that it generates the time evolution of the particle seen by the observer at infinity. The resulting *reduced phase space* \mathcal{P} will, strictly speaking, only be a partially reduced phase space, but it will already be finite dimensional.

With the preparation made in the previous section, the actual phase space reduction is quite simple. The first thing we need is an appropriate set of coordinates on the “constraint surface”, which is defined by (4.1) as a subset of the extended phase space $\tilde{\mathcal{P}}$. This is not very complicated. All we need to do is to take the embedding of spatial infinity (\mathbf{g}, \mathbf{f}) , and extend it to the whole space manifold, such that, everywhere on \mathcal{N} , we have

$$\omega_i = \mathbf{g}^{-1} \partial_i \mathbf{g}, \quad \mathbf{e}_i = \mathbf{g}^{-1} \partial_i \mathbf{f} \mathbf{g}. \quad (4.2)$$

As we already found in section 2, the point particle constraint implies that \mathbf{f} is constant on the particle boundary. There is thus a one-to-one relation between field configurations satisfying the kinematical constraints and embeddings of \mathcal{N} , cut along the radial line at $\varphi = \pm\pi$, with the condition that $\bar{\mathbf{f}} = \mathbf{x}$ is independent of φ .

Using the variables \mathbf{g} and \mathbf{f} as coordinates, we can compute the pullback of the symplectic form Ω on the constraint surface. It is actually simpler to first compute the pullback of the symplectic potential Θ , and then use that $\Omega = d\Theta$. All we need to do is to insert (4.2) into the expression (3.25) for Θ . Thereby, we can use the following trick. We know that the full expression is independent of R , provided that R is large enough, which means that the fall off condition (3.15) must hold. But now, this is fulfilled

for any R . In particular, we can take $R = 0$. Doing so, the first integral vanishes because the integration range shrinks to a set of measure zero. The second integral also vanishes, because $\mathbf{f} = \mathbf{x}$ is constant and therefore $\partial_\varphi \mathbf{f} = 0$ at $r = 0$. Only the last term remains, and there we can replace \mathbf{f}_+ with \mathbf{x} , because this is now also evaluated at $r = 0$. All together, this gives

$$\Theta = -\frac{1}{2}\text{Tr}(\mathbf{u}^{-1}d\mathbf{u}\mathbf{x}). \quad (4.3)$$

The Hamiltonian also takes a rather simple form. As all constraints except the mass shell constraint have been solved, this is the only term that remains,

$$H = \zeta \left(\frac{1}{2}\text{Tr}(\mathbf{u}) - \cos m \right). \quad (4.4)$$

Here, we also used the relation $\bar{\mathbf{u}} = \bar{\mathbf{g}}^{-1}\mathbf{u}\bar{\mathbf{g}}$, to replace $\bar{\mathbf{u}}$ with \mathbf{u} . Thus, the only variables that remain are $\mathbf{x} \in \mathfrak{sl}(2)$ and $\mathbf{u} \in \text{SL}(2)$. These are the only “observables”, that is, the only quantities that are invariant under the gauge transformations generated by the kinematical constraints. We can also reexpress everything in terms of a reduced Lagrangian, which only depends on these variables,

$$L = -\frac{1}{2}\text{Tr}(\mathbf{u}^{-1}\dot{\mathbf{u}}\mathbf{x}) - \zeta \left(\frac{1}{2}\text{Tr}(\mathbf{u}) - \cos m \right). \quad (4.5)$$

To check our calculations, we can rederive the equations of motion from this. Of course, the multiplier ζ still provides the correct mass shell condition. Under a variation of \mathbf{x} , we get

$$\delta L = \frac{1}{2}\text{Tr}(\mathbf{u}^{-1}\dot{\mathbf{u}}\delta\mathbf{x}) \quad \Rightarrow \quad \dot{\mathbf{u}} = 0, \quad (4.6)$$

and varying \mathbf{u} gives

$$\delta L = \frac{1}{2}\text{Tr}(\delta(\mathbf{u}^{-1}\dot{\mathbf{u}})\mathbf{x} + \zeta\delta\mathbf{u}) = \frac{1}{2}\text{Tr}(\mathbf{u}^{-1}\delta\mathbf{u}(\zeta\mathbf{p} - \dot{\mathbf{x}})). \quad (4.7)$$

Here, we used $\delta(\mathbf{u}^{-1}\dot{\mathbf{u}}) = \mathbf{u}^{-1}\partial_t(\delta\mathbf{u}\mathbf{u}^{-1})\mathbf{u}$, and $\dot{\mathbf{u}} = 0$. As a result, we find the evolution equations

$$\dot{\mathbf{u}} = 0, \quad \dot{\mathbf{x}} = \zeta\mathbf{p}, \quad (4.8)$$

which are in agreement with what we found in section 2. In particular, the momentum \mathbf{u} is constant, and the world line of the particle in the embedding space is given by (2.10), with the evolution of the eigentime τ given by the multiplier ζ .

The physical interpretation of the observables \mathbf{x} and \mathbf{u} is now also rather obvious. We found that the value of \mathbf{f} on the particle boundary was the position of the particle in the embedding Minkowski space. This space is now identical with the reference frame of the observer at infinity, and therefore \mathbf{x} is the *position* of the particle, seen by this observer. Similarly, \mathbf{u} , or its projection into the algebra \mathfrak{p} , is the *momentum* of the particle, measured by the same observer. Together they span the reduced phase space $\mathcal{P} = \text{SL}(2) \times \mathfrak{sl}(2)$.

We can now forget about the rather long derivation of reduced phase space \mathcal{P} from Einstein gravity, and summarize the result as follows. The model effectively describes

a point particle from the point of view of an external observer. It has many features in common with a relativistic point particle. The phase space \mathcal{P} is six dimensional, with the basic variables being the position and the momentum of the particle, and the physical subspace of \mathcal{P} is defined by a mass shell constraint. The dynamics of the particle will be described in the same way, namely as a gauge symmetry generated by the mass shell constraint.

There is however one crucial difference. The phase space \mathcal{P} is not a six dimensional vector space, but the product of a group manifold with its Lie algebra, which is the same as the cotangent bundle of the group manifold, $\mathcal{P} = T^*\text{SL}(2)$. Such phase spaces are well known. They typically appear as the phase spaces of, say, particles moving on a curved manifold, or field theories with curved target spaces. But here the situation is somewhat different. It is not the *position* variable that lives on the curved manifold, but the *momentum*, and this will lead to some very uncommon features of this model, especially in the quantum theory.

Poisson brackets

Many of these features can already be seen at the classical level, and most of them follow immediately from the Poisson structure on \mathcal{P} . Exploiting the cotangent space structure, it is more or less straightforward to derive the Poisson bracket. Let us expand the position variable x in terms of the gamma matrices and rewrite the symplectic potential as

$$\Theta = -\frac{1}{2}\text{Tr}(\mathbf{u}^{-1}d\mathbf{u}\gamma_a)x^a. \quad (4.9)$$

The one-form $\frac{1}{2}\text{Tr}(\mathbf{u}^{-1}d\mathbf{u}\gamma_a)$ is the dual of the left invariant vector field \mathcal{L}_a on the group manifold, which is defined by

$$\mathcal{L}_a \mathbf{u} = \frac{1}{2} \mathbf{u} \gamma_a. \quad (4.10)$$

It is therefore reasonable to assume that the action of x^a under the Poisson bracket is associated with this vector field. With this ansatz, and using the Jacobi identity, it is not difficult to derive the basic Poisson brackets, which turn out to be

$$\{\mathbf{u}, x^a\} = \mathbf{u} \gamma^a, \quad \{x^a, x^b\} = 2 \varepsilon^{ab}_c x^c. \quad (4.11)$$

An alternative way to derive this bracket, without using the cotangent space structure, is to consider the components p_a and x^a as independent variables, and insert the expansion $\mathbf{u} = u \mathbf{1} + p^a \gamma_a$, with $u = \sqrt{p^a p_a + 1}$, into the expression for the symplectic potential. It then becomes

$$\Theta = -q^a dp_a, \quad q^a = (\eta^{ab}u + \varepsilon^{abc}p_c - u^{-1}p^a p^b) x_b. \quad (4.12)$$

Hence, it has the canonical form, and the brackets can then be derived from the condition

$$\{p_a, p_b\} = 0, \quad \{p_a, q^b\} = \delta_a^b, \quad \{q^a, q^b\} = 0. \quad (4.13)$$

The result is

$$\{p_a, x^b\} = \delta_a^b u + \varepsilon_a^{bc} p_c, \quad \{u, x^b\} = p^b. \quad (4.14)$$

This also follows from (4.11), by taking the trace or projecting it onto the gamma matrices, and vice versa it implies that x^a acts on \mathbf{u} by right multiplication with γ^a . As a “proof”, we rederive the time evolution equations from the Hamiltonian (4.4),

$$\dot{\mathbf{u}} = \{H, \mathbf{u}\} = 0, \quad \dot{x}^a = \{H, x^a\} = \frac{1}{2}\zeta \text{Tr}(\mathbf{u}\gamma^a) = \zeta p^a. \quad (4.15)$$

As ζ is a free parameter, this is actually a gauge transformation, generated by the mass shell constraint. Once again, this is the same situation as for the relativistic point particle. It is also the reason why we did not divide out all gauge symmetries from the very beginning. We want to keep the remaining gauge freedom when we quantize the model, because we can then easily compare the result to the quantized relativistic point particle.

Before coming to this, we shall however stick to the classical description for a moment and see what the basic *physical* properties of the reduced phase space \mathcal{P} are. The most apparent feature of the Poisson algebra is of course that the components x^a of the position of the particle have non-vanishing brackets. The fact that the corresponding quantum operators do not commute will have profound consequences for the quantum theory. For example, they add another uncertainty relation to the quantum mechanics, which implies that the particle cannot be localized.

We should emphasize that this “non-commutativity of spacetime” is not the result of any choice. The interpretation of the “observable” x as the position of the particle with respect to the observer at infinity is forced upon us, as a consequence of the definitions made at the *field theoretical* level in the beginning. At this point, the description of the particle was completely independent from that of the observer. The only connection between them was the gravitational field, and its dynamics is described by the Einstein Hilbert action. So, the fact that the position coordinates do not commute turns out to be an effect of Einstein gravity.

Symmetries

The last issues we want to investigate at the classical level are the symmetries and their Noether charges. A transformation is a symmetry if the Lagrangian is invariant up to a total time derivative. If this is the case, then the symmetry is realized on phase space as a canonical transformation, and we should find the associated Noether charge that generates the symmetry via the Poisson bracket. As a symmetry also leaves the Hamiltonian invariant, which in this case means that its generator commutes with the mass shell constraint, it represents an observable of the system. The algebra of the Noether charges is therefore simultaneously the observable algebra, which will play a central role in the quantum theory.

Let us first consider Lorentz transformations. It is easily checked that the reduced Lagrangian (4.5) is invariant under

$$\mathbf{u} \mapsto \mathbf{h}^{-1} \mathbf{u} \mathbf{h}, \quad x \mapsto \mathbf{h}^{-1} x \mathbf{h}, \quad (4.16)$$

where $\mathbf{h} \in \text{SL}(2)$ is a constant. At the extended phase space level, this corresponds to a rigid Lorentz transformation of the embedding Minkowski space at spatial infinity. Infinitesimally, the transformation can be written as

$$\delta \mathbf{u} = \frac{1}{2} \lambda^a [\mathbf{u}, \gamma_a], \quad \delta \mathbf{x} = \frac{1}{2} \lambda^a [\mathbf{x}, \gamma_a]. \quad (4.17)$$

To find the Noether charge corresponding to this symmetry, we need a function J_a satisfying

$$\{J_a, \mathbf{u}\} = \frac{1}{2} [\mathbf{u}, \gamma_a], \quad \{J_a, \mathbf{x}\} = \frac{1}{2} [\mathbf{x}, \gamma_a]. \quad (4.18)$$

The action of J_a on \mathbf{u} is the sum of the left and right invariant vector fields, defined by

$$\mathcal{L}_a \mathbf{u} = \frac{1}{2} \mathbf{u} \gamma_a, \quad \mathcal{R}_a \mathbf{u} = -\frac{1}{2} \gamma_a \mathbf{u}. \quad (4.19)$$

They form two commuting copies of an $\mathfrak{so}(1, 2)$ Lie algebra,

$$[\mathcal{L}_a, \mathcal{L}_b] = \varepsilon_{ab}{}^c \mathcal{L}_c, \quad [\mathcal{R}_a, \mathcal{R}_b] = \varepsilon_{ab}{}^c \mathcal{R}_c, \quad [\mathcal{L}_a, \mathcal{R}_b] = 0. \quad (4.20)$$

Moreover, the right invariant vector fields can be written as $\mathcal{R}_a = -u_a{}^b \mathcal{L}_b$, where $u_a{}^b$ is the vector representation of \mathbf{u} (see (A.6)). Using this, it is not difficult to identify the angular momentum as

$$J_a = \frac{1}{2} (u_a{}^b x_b - x_a). \quad (4.21)$$

As expected, its components form an $\mathfrak{so}(1, 2)$ algebra under the Poisson bracket,

$$\{J_a, J_b\} = -\varepsilon_{ab}{}^c J_c. \quad (4.22)$$

If we write it as a matrix, we find that the angular momentum is, up to a factor, the same as the parameter \mathbf{v} appearing in the boundary condition (2.6), which is related to the position and momentum of the particle by (2.8),

$$J^a \gamma_a = \frac{1}{2} (u^a{}_b x^b - x^a) \gamma_a = \frac{1}{2} (\mathbf{u} \mathbf{x} \mathbf{u}^{-1} - \mathbf{x}) = -\frac{1}{2} \mathbf{v}. \quad (4.23)$$

We can also express J_a as a function of the components of the position and the momentum vector. It then becomes

$$J_a = u \varepsilon_{abc} x^b p^c + 2 x_{[a} p_{b]} p^b. \quad (4.24)$$

This expression is different from what one would expect naively, namely $\varepsilon_{abc} x^b p^c$. Only in the “low momentum” limit, where p_a is small compared to $u \approx 1$, it reduces to the standard expression.

If we look at the action of J_a and x_a on \mathbf{u} , we find that there is a remarkable similarity. If $F(\mathbf{u})$ is some function on $\text{SL}(2)$, then we have

$$\{x_a, F\} = -2 \mathcal{L}_a F, \quad \{J_a, F\} = (\mathcal{L}_a + \mathcal{R}_a) F. \quad (4.25)$$

The left and right invariant vector fields \mathcal{L}_a and \mathcal{R}_a span an $\mathfrak{so}(1, 2) \times \mathfrak{so}(1, 2) \simeq \mathfrak{so}(2, 2)$ Lie algebra. This is the symmetry algebra of anti-de-Sitter space. Indeed,

at least locally, our momentum space looks like anti-de-Sitter space. If we use the coordinates $p_A = (p_a, u)$, then they obey

$$p_A p_B \eta^{AB} = -1, \quad \eta^{AB} = \text{diag}(-1, 1, 1, -1). \quad (4.26)$$

Hence, it is the unit hyperboloid in $\mathbb{R}^{2,2}$, and its isometries are generated by the vector fields

$$\mathcal{G}_{AB} = \frac{1}{2} \left(p_A \frac{\partial}{\partial p^B} - p_B \frac{\partial}{\partial p^A} \right). \quad (4.27)$$

If we split them into

$$\mathcal{G}_{ab} = \frac{1}{2} \left(p_a \frac{\partial}{\partial p^b} - p_b \frac{\partial}{\partial p^a} \right), \quad \mathcal{G}_a = \frac{1}{2} \left(u \frac{\partial}{\partial p^a} + p_a \frac{\partial}{\partial u} \right), \quad (4.28)$$

then it is not difficult to verify that

$$\{x_a, F\} = -2 \mathcal{G}_a F - \varepsilon_a{}^{bc} \mathcal{G}_{bc} F, \quad \{J_a, F\} = \varepsilon_a{}^{bc} \mathcal{G}_{bc} F, \quad (4.29)$$

where $F(u)$ is considered as a function of the components p_A . Hence, x_a and J_a generate the isometries of momentum space. This is again analogous to the relativistic point particle. There, the momentum space is Minkowski space, its isometry group is $\text{ISO}(1, 2)$, and the components of the position and angular momentum vectors are the generators of translations and Lorentz transformations. Here, the situation is exactly the same, except that the isometry group of momentum space is $\text{SO}(2, 2)$. This is the reason why the algebra of the position and angular momentum components is different.

The idea of momentum space being anti-de-Sitter is actually not new. It was proposed in 1946(!) by Snyder in the context of relativistic point particles in 3+1 dimensions [12]. He realized that this implies that the components of the position vector of a particle can no longer commute with each other. Quite remarkably, it also introduces a shortest distance into the quantum theory, without breaking the Lorentz symmetry, as it typically happens when a lattice structure is introduced on spacetime. In other words, it is possible to introduce a cutoff in spacetime in a covariant way. At the quantum level, we shall find that our model has all these features as well.

So far, we considered the isometries of the *momentum* space. But what about those of the *position* space, or the *spacetime* in which the particle lives. As this is flat Minkowski space, we should expect that there is an $\text{ISO}(1, 2)$ symmetry group, and using the analogy to the relativistic point particle, the symmetries should be generated by the components of the momentum p_a and the angular momentum J_a . Indeed, these quantities form a Poincaré algebra,

$$\{p_a, p_b\} = 0, \quad \{J_a, p_b\} = -\varepsilon_{ab}{}^c p_c, \quad \{J_a, J_b\} = -\varepsilon_{ab}{}^c J_c. \quad (4.30)$$

We know already that J_a is the Noether charge associated with Lorentz transformations (4.18). So, we should expect that p_a is the generator of a translation of the particle. Let us contract the momentum with an arbitrary vector n^a , and compute the bracket with the position coordinate. What we get is

$$\{n^b p_b, x^a\} = u n^a - \varepsilon^{abc} n_b p_c. \quad (4.31)$$

As the p 's themselves have vanishing brackets with each other, we can easily exponentiate this transformation, and write down the action of the symmetry on the basic phase space variables, parametrized by a vector $\mathbf{n} = n^a \gamma_a$,

$$\mathbf{u} \mapsto \mathbf{u}, \quad \mathbf{x} \mapsto \mathbf{x} + u \mathbf{n} + [\mathbf{n}, \mathbf{p}]. \quad (4.32)$$

This does not look like an ordinary translation $\mathbf{x} \mapsto \mathbf{x} + \mathbf{n}$. Maybe the reason for this is that the components of the momentum \mathbf{p} are just not the correct generators. To see that this is not the case, consider the transformation of the Lagrangian (4.5) under an ordinary translation. It is given by

$$L \mapsto L - \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \dot{\mathbf{u}} \mathbf{n}), \quad (4.33)$$

and this is not a total time derivative. So, the ordinary translations are *not* symmetries of the action. One can also easily see that the Poisson bracket is not invariant under ordinary translations. But the deformed translations (4.32) are canonical transformations, and the Lagrangian is invariant. Only in the low momentum limit, where we can neglect the last term in (4.32) and set $u \approx 1$, the deformed translations tend to the usual ones.

Let us close the classical part with this somewhat peculiar observation. The model that we derived is, for large momenta, not invariant under ordinary translations, although the position of the particle lives in flat Minkowski space. There is however, and maybe this is even more remarkable, still a Poincaré group of canonical transformations acting on the phase space, which is generated by the components of the momentum and angular momentum. It is thus not the isometry group itself that is “deformed”, but only its action on the phase space variables. Under a translation, the particle is not shifted into the direction of the translation, but it receives an additional shift into a direction determined by its momentum. At the quantum level, this absence of the ordinary translational symmetry will show up even more drastically.

5 Quantum theory

We now want to go over to a quantum description of the point particle, based on the classical reduced phase space formulation. We shall first consider the quantization of the phase space \mathcal{P} itself, without imposing the mass shell condition. The philosophy behind this is that we would first like to study the quantum spacetime on which the particle lives. Typical questions that we would like to answer are:

- What is the spectrum of the operators x_a , p_a , and J_a ?
- How are the uncertainty relations modified?
- Does a shortest length scale show up in the model?

Later we will consider the dynamics of the particle by imposing the mass shell constraint, which translates into a kind of Klein Gordon equation, or even a Schrödinger like time evolution equation.

We have seen in the previous sections that the phase space \mathcal{P} is the cotangent bundle $T^*\text{SL}(2)$, or $\text{SL}(2) \times \mathfrak{sl}(2)$. A point in that space was represented by a pair (\mathbf{u}, \mathbf{x}) , where $\mathbf{u} \in \text{SL}(2)$ is the momentum and $\mathbf{x} \in \mathfrak{sl}(2)$ the position of the particle. The most remarkable feature of the Poisson structure on \mathcal{P} was that the components of \mathbf{x} , and hence the coordinates of the particle, do not commute. This implies that we cannot represent quantum states as wave functions in spacetime. We can however set up a momentum representation in which the wave function depends on \mathbf{u} .

Momentum representation

From the mathematical point of view, this is the most natural way to quantize the given phase space, which is the cotangent bundle of the momentum space. Hence, we take as our Hilbert space $\mathcal{H} = \mathcal{L}_2(\text{SL}(2), d\mathbf{u})$, where $d\mathbf{u}$ is the Haar measure. A state $|\Psi\rangle \in \mathcal{H}$ is then represented by a wave function $\Psi(\mathbf{u})$. Using the bra-ket notation, we can write $\Psi(\mathbf{u}) = \langle \mathbf{u} | \Psi \rangle$, where $|\mathbf{u}\rangle$ are the momentum eigenstates, $\hat{\mathbf{u}}|\mathbf{u}\rangle = \mathbf{u}|\mathbf{u}\rangle$. They are orthonormal and complete,

$$\langle \mathbf{u}_1 | \mathbf{u}_2 \rangle = \delta(\mathbf{u}_1^{-1} \mathbf{u}_2), \quad \int d\mathbf{u} |\mathbf{u}\rangle \langle \mathbf{u}| = \hat{1}. \quad (5.1)$$

Using this, we can write the scalar product and the expansion in terms of a wave function as

$$\langle \Psi | \Phi \rangle = \int d\mathbf{u} \Psi^*(\mathbf{u}) \Phi(\mathbf{u}), \quad |\Psi\rangle = \int d\mathbf{u} \Psi(\mathbf{u}) |\mathbf{u}\rangle. \quad (5.2)$$

To define an operator representation for phase space functions, we demand that real functions should be represented by Hermitian operators, and that for the basic variables we have

$$\{A, B\} = C \quad \Rightarrow \quad [\hat{A}, \hat{B}] = -i\hbar \hat{C}. \quad (5.3)$$

In particular, we want that

$$\{x_a, \mathbf{u}\} = -\mathbf{u}\gamma_a \quad \Rightarrow \quad [\hat{x}_a, \hat{\mathbf{u}}] = i\hbar \hat{\mathbf{u}}\gamma_a. \quad (5.4)$$

From this we infer that x_a has to act on the basis states as

$$\hat{x}_a |\mathbf{u}\rangle = -2i\hbar \mathcal{L}_a |\mathbf{u}\rangle, \quad (5.5)$$

where the \mathcal{L}^a are the left invariant vector fields acting on \mathbf{u} , as defined in (4.19). The Haar measure is invariant under the action of the left and right invariant vector fields, so that the operator x_a is Hermitian. Similarly, we can derive the operator for the angular momentum from

$$\{J_a, \mathbf{u}\} = \frac{1}{2}[\mathbf{u}, \gamma_a] \quad \Rightarrow \quad [\hat{J}_a, \hat{\mathbf{u}}] = -\frac{1}{2}i\hbar (\hat{\mathbf{u}}\gamma_a - \gamma_a\hat{\mathbf{u}}), \quad (5.6)$$

which gives

$$\hat{J}_a |\mathbf{u}\rangle = i\hbar (\mathcal{L}_a + \mathcal{R}_a) |\mathbf{u}\rangle. \quad (5.7)$$

Clearly, this is again Hermitian. We can express it in terms of x^a and \mathbf{u} as

$$i\hbar (\mathcal{L}_a + \mathcal{R}_a) |\mathbf{u}\rangle = i\hbar (\mathcal{L}_a - u_a^b \mathcal{L}_b) |\mathbf{u}\rangle = \frac{1}{2} (\hat{u}_a^b \hat{x}_b - \hat{x}_a) |\mathbf{u}\rangle, \quad (5.8)$$

which is the same as the classical expression (4.21). There are no ordering ambiguities, because $[\hat{u}_a^b, \hat{x}_b] = 0$.

Euler angle representation

To derive the spectrum of the operators x_a , p_a , and J_a , we have to make things a bit more explicit. For this purpose, we switch to a slightly different representation. Actually, we only change the naming of the basis states. Instead of using the matrix \mathbf{u} to label the state $|\mathbf{u}\rangle$, we use the Euler angles as coordinates on the group manifold. The state $|\rho, \chi, \phi\rangle$ is then the same as $\ker \mathbf{u}$, where

$$\mathbf{u} = e^{\frac{1}{2}(\rho+\phi)\gamma_0} e^{\chi\gamma_1} e^{\frac{1}{2}(\rho-\phi)\gamma_0}. \quad (5.9)$$

The periodicity of the coordinates can then be expressed in the relations

$$|\rho, \chi, \phi\rangle = |\rho + 2\pi, \chi, \phi\rangle = |\rho, \chi, \phi + 2\pi\rangle. \quad (5.10)$$

To write down the scalar product and the wave function expansion, we need the Haar measure in terms of these coordinates, which is given by

$$d\mathbf{u} = \frac{1}{2\pi^2} \sinh(2\chi) d\rho d\chi d\phi \quad (5.11)$$

Hence, we have

$$|\Psi\rangle = \frac{1}{2\pi^2} \int \sinh(2\chi) d\rho d\chi d\phi \Psi(\rho, \chi, \phi) |\rho, \chi, \phi\rangle, \quad (5.12)$$

and the scalar products of two such states is given by

$$\langle \Phi | \Psi \rangle = \frac{1}{2\pi^2} \int \sinh(2\chi) d\rho d\chi d\phi \Phi^*(\rho, \chi, \phi) \Psi(\rho, \chi, \phi). \quad (5.13)$$

Of course, the states $|\rho, \chi, \phi\rangle$ are still momentum eigenstates, and the eigenvalues are given by

$$\begin{aligned} \hat{u} &\rightarrow \cosh \chi \cos \rho, & \hat{p}_1 &\rightarrow \sinh \chi \cos \phi, \\ \hat{p}^0 &\rightarrow \cosh \chi \sin \rho, & \hat{p}_2 &\rightarrow \sinh \chi \sin \phi. \end{aligned} \quad (5.14)$$

Time and angular momentum eigenstates

Now, let us consider the 0-components of the position and the angular momentum operator. Physically, these correspond to the time $T = x^0$ and the spatial angular momentum $M = J^0$, which is the momentum associated with spatial rotations. To find their action on the basis states, we can use that

$$\frac{\partial \mathbf{u}}{\partial \rho} = (\mathcal{R}^0 - \mathcal{L}^0) \mathbf{u}, \quad \frac{\partial \mathbf{u}}{\partial \phi} = (\mathcal{R}^0 + \mathcal{L}^0) \mathbf{u}. \quad (5.15)$$

From this we infer that

$$\widehat{T}|\rho, \chi, \phi\rangle = i\hbar \left(\frac{\partial}{\partial \rho} - \frac{\partial}{\partial \phi} \right) |\rho, \chi, \phi\rangle, \quad \widehat{M}|\rho, \chi, \phi\rangle = i\hbar \frac{\partial}{\partial \phi} |\rho, \chi, \phi\rangle. \quad (5.16)$$

The eigenstates of these operators are then easily found to be

$$|t, n; \psi\rangle = \frac{1}{2\pi^2} \int \sinh(2\chi) d\rho d\chi d\phi e^{i(t+n)\rho} e^{in\phi} \psi(\chi) |\rho, \chi, \phi\rangle, \quad (5.17)$$

where ψ is an arbitrary function that depends on χ only. The eigenvalues are

$$\widehat{T}|t, n; \psi\rangle = t\hbar |t, n; \psi\rangle, \quad \widehat{M}|t, n; \psi\rangle = n\hbar |t, n; \psi\rangle. \quad (5.18)$$

Both quantum numbers t and n have to be integers, because both ρ and ϕ have a period of 2π . For the spatial angular momentum M , this is of course what we should expect. It is quantized in units of \hbar . The quantization of the eigenvalues of T is however more interesting. It implies that time is quantized. It only takes values which are integer multiples of \hbar . In our units with $G = 1/4\pi$, \hbar also has the dimension of a length or time. We define this to be the Planck length, $\ell_P = 4\pi G\hbar$, and use the symbol ℓ_P in the following instead of \hbar whenever it appears as a length or time scale.

Besides the discreteness, there is another peculiar feature of the spectrum of T . It starts at $t = 0$, and not, say, at $t = \frac{1}{2}$ or $t = 0.62$. What is so special about the integer values of t ? The answer is that there is actually nothing special. The origin of the spectrum is due to an ambiguity in the definition of the Hilbert space, or rather the operator representation introduced above, which we ignored so far. The classical phase space is not simply connected. As a result, there is a phase ambiguity in the definition of the basis states $|\rho, \chi, \phi\rangle$. They have to be periodic in ρ . However, instead of (5.10), it is sufficient to require

$$|\rho + 2\pi, \chi, \phi\rangle = e^{2\pi i\tau} |\rho, \chi, \phi\rangle, \quad (5.19)$$

for some real number τ . Doing so, it is not difficult to see that the spectrum of T is shifted and becomes $(\tau + \mathbb{Z})\ell_P$, instead of $\mathbb{Z}\ell_P$. We can shift it wherever we like by choosing a suitable “quantization parameter” τ . But the time steps will always have the same size ℓ_P . For simplicity, we shall restrict to $\tau = 0$. Although the details of the “spectrum of spacetime” to be derived in the following depend on τ , the *qualitative* behaviour will always be the same, and this is what we are mainly interested in.

The spectrum of spacetime

To derive something like a position representation for the particle, we should try to diagonalize as many x -variables as possible. We cannot diagonalize all three components, because they do not commute. But we can at least achieve to find eigenstates of two functions thereof. In fact, there is an obvious second function that commutes with $T = x^0$, and can be diagonalized simultaneously. As the position operators obey an algebra that is similar to the Euclidean angular momentum algebra, it is not surprising that this is the “Casimir” $S^2 = x^a \hat{x}_a$. It is invariant under Lorentz transformations, so

it also commutes with $M = J^0$, and we can stick to the quantum numbers t and n . On the \mathbf{u} -eigenstates, the operator S^2 acts as

$$\widehat{S}^2 |\mathbf{u}\rangle = -4 \ell_{\text{P}}^2 \mathcal{L}_a \mathcal{L}^a |\mathbf{u}\rangle = -4 \ell_{\text{P}}^2 \Delta |\mathbf{u}\rangle, \quad (5.20)$$

where Δ is the Laplace Beltrami operator, which is equal to the square of the left (or right) invariant vector fields. This leads us to the harmonic analysis on $\text{SL}(2)$. Without going too deep into the mathematical details, we give a brief derivation of the basic results in Appendix B. They are summarized as follows. The eigenstates can be split into two classes, the continuous and the discrete series. For the continuous series, we have states labeled by a positive real quantum number λ , and the two integers t and n , such that

$$\widehat{S}^2 |t, n, \lambda\rangle = (\lambda^2 + 1) \ell_{\text{P}}^2 |t, n, \lambda\rangle. \quad (5.21)$$

The discrete series is labeled by another integer l , and the eigenvalues are given by

$$\widehat{S}^2 |t, n, l\rangle = -l(l-2) \ell_{\text{P}}^2 |t, n, l\rangle. \quad (5.22)$$

Here, the range of the quantum numbers t , n and l are restricted as follows. For a fixed t , the positive integer l only takes finitely many values, namely

$$2 \leq l \leq |t|, \quad l \equiv t \pmod{2}. \quad (5.23)$$

Hence, for even t we have $l = 2, 4, 6, \dots, |t|$, and $l = 3, 5, 7, \dots, |t|$ for odd t . There are no discrete states for t equal to $-1, 0$, or 1 . The values for n are also restricted. For given t and l , we must have

$$n \geq -\frac{1}{2}(t-l) \quad \text{for } t \geq 2, \quad n \leq \frac{1}{2}(t-l) \quad \text{for } t \leq -2. \quad (5.24)$$

Now, observe that the continuous series has positive eigenvalues of S^2 , starting at $S_{\text{min}}^2 = \ell_{\text{P}}^2$. This corresponds to *spacelike* vectors x^a . The discrete series has zero and negative eigenvalues and corresponds to *lightlike* or *timelike* vectors x^a . We can illustrate the situation as follows. In a commutative spacetime, we could set up a position representation of the Hilbert space, in which each point in spacetime corresponds to one basis state. Here, we can no longer assign a basis state to every point in spacetime. We can only fix the values of $S^2 = x^a x_a$ and $T = x^0$. Alternatively, we may also use the quantity $R^2 = (x_1)^2 + (x_2)^2 = S^2 + T^2$ instead of S^2 . Its spectrum is strictly positive,

$$\begin{aligned} \widehat{R}^2 |t, n, \lambda\rangle &= (t^2 + \lambda^2 + 1) \ell_{\text{P}}^2 |t, n, \lambda\rangle, \\ \widehat{R}^2 |t, n, l\rangle &= (t^2 - l(l-2)) \ell_{\text{P}}^2 |t, n, l\rangle, \end{aligned} \quad (5.25)$$

and a fixed value of R^2 defines a *circle* in space. Hence, we cannot associate states with points in spacetime, but at least with circles, labeled by the time coordinate T and the spatial radius squared R^2 . This reflects the fact that a state cannot be localized in spacetime, due to the commutator relation and thus the uncertainty relation between the coordinates.

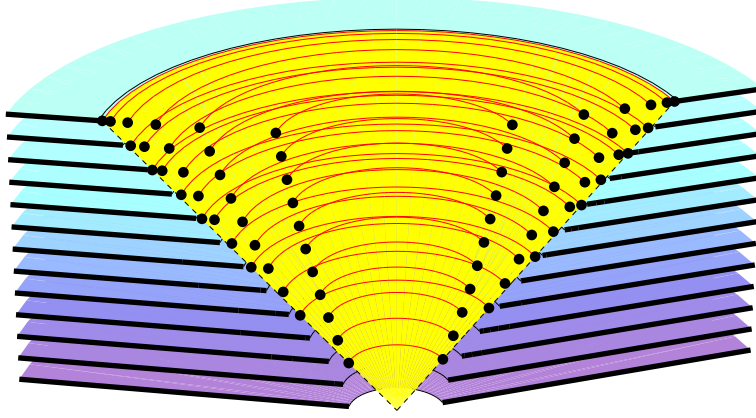


Figure 3: The spectrum of quantum spacetime. Each circle in this diagram corresponds to a tower of states labeled by the spatial angular momentum quantum number n . The position of the circles in Minkowski space are determined by the quantum numbers t and s^2 .

A picture of this quantum spacetime is shown in figure 3. It shows a circle for each allowed value of R^2 and T , for $0 \leq T \leq 12\ell_P$. We get discrete layers of states, corresponding to the discrete eigenvalues of T . Outside the light cone, we have the continuous spectrum of R^2 , starting at $R_{\min}^2 = T^2 + \ell_P^2$, a little bit off the light cone. On the light cone, we have the discrete lightlike states for $|T| \geq 2\ell_P$, and inside the light cone the discrete timelike states which exist for $|T| \geq 3\ell_P$. Each ring in this picture corresponds to an infinite tower of states, labeled by the angular momentum quantum number n . In the spacelike region, where the continuous set of rings forms a plane, any value of n is allowed. On the light cone and in the timelike region, n has a lower limit. On the innermost rings, where $l = t$, n must be positive or zero. The further we go away from the time axis towards the light cone, the more negative values are allowed for n . In the past light cone, the situation is just the reverse. There, only negative n are allowed on the innermost rings, and more and more positive ones are mixed in if we move towards the light cone.

This semi discrete structure of spacetime is a consequence of the fact that the coordinates x^a do not commute. In this sense, figure 3 can be thought of as a picture of a non-commutative spacetime. There is also a new kind of uncertainty relation in the theory. It is impossible to determine any two of the three x^a simultaneously with infinite precision. For example, we have

$$[\hat{x}_1, \hat{x}_2] = 2i\ell_P \hat{x}^0 = 2i\ell_P \hat{T} \Rightarrow \langle \Delta x_1 \rangle \langle \Delta x_2 \rangle \geq 2\ell_P \langle T \rangle. \quad (5.26)$$

This tells us that it is not possible to find a state that is localized inside a spacelike circle with a radius smaller than $\sqrt{2\ell_P T}$. This is exactly the behaviour of the innermost rings in the timelike region of figure 3. The radius of the ring that is closest to the time

axis, and therefore represents the “most localized” state at the time T , grows with the square root of T . Note, however, that the only completely localized state that would be allowed by the uncertainty relation above, namely at the “origin of spacetime” $x = 0$, does not exist. For low values of T the uncertainty is therefore even stronger.

The origin and orientation of spacetime

We already saw at the classical level that the theory is not invariant under ordinary translations $x \mapsto x + n$. At the quantum level, this can be seen in the spectrum of spacetime, which is far from being invariant under translations. The origin seems to play a special role, in the sense that it is the light cone emerging from *this* point that splits the spectrum into the discrete and the continuous series. The picture is also not invariant under Lorentz transformations. However, this is only due to our choice of a special time axis and not of a physical nature. We have the same situation for the angular momentum eigenstates in three dimensional Euclidean space. There one also has to pick out a special z -axis to introduce the l and m quantum numbers, and this brakes the rotational invariance. However, the origin of angular momentum space is a well defined physical concept. It corresponds to a system that does not rotate. But what is the physical meaning of the special point in spacetime?

First of all, we should note that this special point is a *parameter* that enters the action. Implicitly, it shows up in the definition of the boundary term (3.21), which we added to the Lagrangian to render it finite. This term was actually determined up to a constant only, that is, up to a term that does not depend on the regularization parameter R . There is not much choice for such a term, if we demand that it is linear in the velocities, that the Lorentz symmetry is not broken, and that the equations of motion are not modified. What we then still can do is to replace the embedding field f by $f - n$. It results in a redefinition of the Lagrangian,

$$L \mapsto L + \frac{1}{2} \text{Tr}(u^{-1} \dot{u} n). \quad (5.27)$$

This does *not* change the classical equations of motion. It *does* however change the Poisson structure on the phase space and thus the algebra of the quantum operators. The consequence is that the origin of the spectrum shown in the figure is shifted to the point n . We can also see that the transformation (5.27) exactly cancels (4.33), which resulted from an ordinary translation of the position coordinates. The ordinary translations are, in a sense, represented as a kind of “metasymmetry”. The Lagrangian itself is not invariant, but it contains a parameter, and changing this parameter compensates for the translation. We should view upon this as a parameter of the theory, to be determined by “experiments”.

So, we can shift this origin of spacetime wherever we like, but we cannot avoid that there is such a special point. On the other hand, we saw that, even with the origin fixed, there is a perfectly well defined action of the Poincaré group on the classical phase space \mathcal{P} , realized by the Lorentz rotations and the deformed translations. Of course, these symmetries are also realized as unitary transformations on \mathcal{H} , generated by the operators J_a and p_a . However, the deformed translations mix the position and momentum variables in a somewhat strange way, so that the result is not a simple shift of the spectrum of spacetime. One could interpret this as a “deformed principle of

relativity”. If one observer sees a classical state (\mathbf{u}, \mathbf{x}) , or a quantum state $|\Psi\rangle$, then another observer, which is located at a different point in spacetime, sees a different state, and this state is not obtained by a simple shift in \mathbf{x} but by a deformed translations (4.32), or its quantum analogue. This would restore the translational invariance, and it also provides a physical interpretation of the origin, which becomes the position of the observer.

Of course, all these statements are somewhat counterintuitive. A different way out of this dilemma might be to consider multi particle models. For a system of finitely many particles, it turns out that the Poisson structure on the reduced phase space is very similar to (4.11), except that now the variables \mathbf{x} and \mathbf{u} become the *relative* coordinates. They represent the relative distance between two particles, and something like their relative momentum [13]. As a result, one finds that the spectrum of the relative coordinates looks like the one shown in figure 3. The origin then has a well defined meaning. It corresponds to a state where the two particles are at the same point in spacetime. The fact that there is no such state in the quantum spectrum, and that for the state that is closest to the origin we have $S_{\min}^2 = \ell_P^2$, implies that the particles cannot come closer to each other than one Planck length.

Another issue, which is also due to an ambiguity in the definition of the boundary term at infinity, concerns the inherent orientation of spacetime. At the classical level, this is already apparent from the basic brackets between the coordinates x^a . The appearance of the Levi Civita symbol indicates that the theory is not “ PT ” invariant. The Lagrangian is not invariant with respect to the substitutions $\mathbf{u} \mapsto \mathbf{u}^{-1}$ (or $\mathbf{p} \mapsto -\mathbf{p}$) and $\mathbf{x} \mapsto -\mathbf{x}$. Again, this broken symmetry can be seen quite nicely in the quantum spacetime picture. On the discrete rings in the future light cone which are closest the time axis, only positive values of the spatial angular momentum occur. The further we go away from the time axis, the more negative angular momenta mix in. In the past light cone, the situation is reversed. Slightly sloppy speaking, if we want to localize the particle on such a ring, then we have to give it an angular momentum in a certain direction.

This orientation results from another freedom in the definition of the boundary term (3.21). We can replace \mathbf{f}_+ by a linear combination of \mathbf{f}_+ and \mathbf{f}_- , which leads to a more general reduced Lagrangian, whose kinetic part is parametrized by a real number σ ,

$$L_{\text{kin}} = -\frac{1}{2}\sigma \text{Tr}(\dot{\mathbf{u}} \mathbf{u}^{-1} \mathbf{x}) - \frac{1}{2}(1 - \sigma) \text{Tr}(\mathbf{u}^{-1} \dot{\mathbf{u}} \mathbf{x}). \quad (5.28)$$

Again, this does not change the equations of motion, but it does affect the Poisson structure. Replacing σ with $1 - \sigma$ corresponds to a “ PT ” transformation $\mathbf{x} \mapsto -\mathbf{x}$, $\mathbf{u} \mapsto \mathbf{u}^{-1}$, so that this symmetry is also realized as a kind of metasymmetry. Unfortunately, the value $\sigma = \frac{1}{2}$, which would remove the handedness of spacetime, is the only value for σ for which the classical equations of motion are no longer reproduced fully. This can be seen by repeating the derivation of the evolution equations (4.8) with the modified Lagrangian. So, we have to take the same point of view here as well. The orientation of spacetime is a parameter which has to be determined by experiments. For simplicity, let us stick to $\sigma = 0$ in the following.

The physical Hilbert space

Let us now consider the dynamics of the particle on the quantum spacetime, which is described by the mass shell constraint. Classically, it was given by

$$C = u - \cos m = 0. \quad (5.29)$$

Let us see how the corresponding operator acts in the Euler angle representation. We find that

$$\begin{aligned} \hat{C} |\rho, \chi, \phi\rangle &= (\hat{u} - \cos m) |\rho, \chi, \phi\rangle \\ &= (\cos \rho \cosh \chi - \cos m) |\rho, \chi, \phi\rangle. \end{aligned} \quad (5.30)$$

If we expand a generic state as in (5.12),

$$|\Psi\rangle = \frac{1}{2\pi^2} \int \sinh(2\chi) d\rho d\chi d\phi \Psi(\rho, \chi, \phi) |\rho, \chi, \phi\rangle, \quad (5.31)$$

then the wave function has to satisfy

$$(\cos \rho \cosh \chi - \cos m) \Psi(\rho, \chi, \phi) = 0. \quad (5.32)$$

It states that Ψ has support on the mass shells only, and defines the physical state space $\mathcal{H}_{\text{phys}}$ as a subset of \mathcal{H} . We can use ϕ and χ as coordinates on the mass shells, and write the general solution to the mass shell constraint as

$$\Psi(\rho, \chi, \phi) = \sum_{\epsilon} \delta(\cos \rho \cosh \chi - \cos m) \theta(\epsilon \sin \rho) \Psi_{\epsilon}(\chi, \phi). \quad (5.33)$$

Here, $\Psi_{\epsilon}(\chi, \phi)$, with $\epsilon = \pm 1$, is a pair of arbitrary functions. We have $\epsilon = 1$ for particles and $\epsilon = -1$ for antiparticles. The theta function, which is 0 for negative arguments and 1 for positive, ensures that $0 < \rho < \pi$ for particles and $-\pi < \rho < 0$ for antiparticles.

As for the relativistic point particle, we are faced with the problem that the physical state space $\mathcal{H}_{\text{phys}}$ is actually not a subspace of the Hilbert space \mathcal{H} . The wave functions (5.33) are not normalizable. The problem can be solved in the standard way, by introducing a new scalar product, which is, up to a constant, uniquely determined if we require it to be Lorentz invariant. In other words, we want the operators J_a and p_a to remain Hermitian under the new product.

To find this physical product explicitly, we note that the wave function $\Psi_{\epsilon}(\chi, \phi)$ in (5.33) transforms as a scalar under Lorentz transformations. The left hand side is a scalar function of $\mathbf{u} = \mathbf{u}(\rho, \chi, \phi)$, and the argument of the delta function is also a scalar, because it is equal to $u - \cos m$. Hence, to define a Lorentz invariant scalar product, we need a Lorentz invariant measure on the mass shell in terms of the coordinates χ and ϕ . This is found to be

$$\langle \Phi | \Psi \rangle_{\text{phys}} = \frac{1}{4\pi^2} \sum_{\epsilon} \int \frac{2 \sinh(2\chi) d\chi d\phi}{\sqrt{\sinh^2 \chi + \sin^2 m}} \Phi_{\epsilon}^*(\chi, \phi) \Psi_{\epsilon}(\chi, \phi). \quad (5.34)$$

It becomes a more familiar expression, if we replace the polar coordinates (χ, ϕ) by the components of the spatial momentum vector $\vec{p} = (p_1, p_2)$. Explicitly, we have

$$p_1 = \sinh \chi \cos \phi, \quad p_2 = \sinh \chi \sin \phi, \quad (5.35)$$

and this implies

$$\frac{2 \sinh(2\chi) d\chi d\phi}{\sqrt{\sinh^2 \chi + \sin^2 m}} = \frac{d^2 \vec{p}}{\sqrt{p^2 + \sin^2 m}} = \frac{d^2 \vec{p}}{\omega(\vec{p})}. \quad (5.36)$$

On the mass shell, $\omega(\vec{p})$ is equal to $|p^0|$, so that the expression for the scalar product becomes formally the same as for the relativistic point particle in the momentum representation,

$$\langle \Phi | \Psi \rangle_{\text{phys}} = \frac{1}{4\pi^2} \sum_{\epsilon} \int \frac{d^2 \vec{p}}{\omega(\vec{p})} \Phi_{\epsilon}^*(\vec{p}) \Psi_{\epsilon}(\vec{p}). \quad (5.37)$$

The Klein Gordon equation

As a next step, we would like to see in which sense the mass shell constraint can be interpreted as a time evolution equation for the quantum state. To explain what we are going to do, let us consider a relativistic point particle in flat spacetime for a moment. There we have the mass shell constraint in momentum space,

$$(p_a p^a + m^2) \Psi(p) = 0. \quad (5.38)$$

Clearly, to see that this is actually a time evolution equation, we only need to Fourier transform to the position representation. But what if there was no such representation, as in our model? We can then still see that it is a time evolution equation, if we diagonalize another set of operators, such that one of them is the time operator $T = x^0$. Let us choose as the other two operators the absolute length of the spatial momentum $p^2 = p_1^2 + p_2^2$, and the spatial angular momentum $M = x_1 p_2 - x_2 p_1$. They commute with each other and both commute with T . If we now write the transformed wave function as $\Psi(t, n; p)$, where $t \in \mathbb{R}$, $n \in \mathbb{Z}$, and $p \in \mathbb{R}_+$, then the mass shell constraint becomes a second order differential equation in t ,

$$-\hbar^2 \frac{\partial^2}{\partial t^2} \Psi(t, n; p) = (p^2 + m^2) \Psi(t, n; p). \quad (5.39)$$

So, without transforming completely to the position representation, we find this ‘‘Klein Gordon equation’’. We can convert it into a Schrödinger equation, if we decompose the solutions into particle and antiparticle states. The general solution to (5.39) can then be written as

$$\Psi(t, n; p) = \sum_{\epsilon} \Psi_{\epsilon}(t, n; p), \quad (5.40)$$

where the two functions Ψ_{ϵ} , with $\epsilon = \pm 1$, have to obey

$$-i\hbar \frac{\partial}{\partial t} \Psi_{\epsilon}(t, n; p) = \epsilon \sqrt{p^2 + m^2} \Psi_{\epsilon}(t, n; p). \quad (5.41)$$

We chose this somewhat skew representation, because in our model we can diagonalize exactly the same operators. We did this already in (5.17), which was the most general eigenstate of T and M . Let us define

$$|t, n; \chi\rangle = \frac{1}{4\pi^2} \int d\rho d\phi e^{i(t+n)\rho} e^{in\phi} |\rho, \chi, \phi\rangle, \quad (5.42)$$

and expand the state in this basis,

$$|\Psi\rangle = \sum_{t \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \int d\chi 2 \sinh(2\chi) \Psi(t, n; \chi) |t, n; \chi\rangle. \quad (5.43)$$

As the rapidity χ is related to the spatial momentum by $p = \sinh \chi$, this is essentially the same representation as the one considered above for the relativistic point particle. To write the mass shell constraint as an equation for the wave function $\Psi(t, n; \chi)$, we need to know how u acts on the new basis states,

$$\begin{aligned} \hat{u} |t, n; \chi\rangle &= \frac{1}{4\pi^2} \int d\rho d\phi e^{i(t+n)\rho} e^{in\phi} \cos \rho \cosh \chi |\rho, \chi, \phi\rangle \\ &= \frac{1}{2} \cosh \chi (|t+1, n; \chi\rangle + |t-1, n; \chi\rangle). \end{aligned} \quad (5.44)$$

Using this, we can write the Klein Gordon equation as

$$\Psi(t+1, n; \chi) + \Psi(t-1, n; \chi) = 2 \frac{\cos m}{\cosh \chi} \Psi(t, n; \chi). \quad (5.45)$$

This is not a second order *differential* equation, but a second order *difference* equation. We should have expected something like this, because the eigenvalues of T are discrete. The fact that it is such a nice nearest neighbour equation is a result of the appearance of $\cos \rho$ in the constraint. It is not difficult to see that it reduces to the ordinary Klein Gordon equation in the limit $\ell_P \rightarrow 0$. To do this, we set $\sinh \chi = p$, and then make the replacements

$$m \mapsto 4\pi G m, \quad p \mapsto 4\pi G p, \quad t \mapsto t/\ell_P, \quad (5.46)$$

to restore the physical units. If we then expand (5.45) for small ℓ_P , but finite \hbar , and use that $\ell_P = 4\pi G \hbar$, we recover (5.39) in the leading order, which is ℓ_P^2 .

We can also convert the discretized Klein Gordon equation into a Schrödinger equation, if we split the wave function into a particle and an antiparticle wave function. This becomes more or less straightforward, if we first write the step equation as

$$\Psi(t+1, n; \chi) + \Psi(t-1, n; \chi) = (U + U^{-1}) \Psi(t, n; \chi), \quad (5.47)$$

where U is found to be

$$U = \frac{\cos m + i \sqrt{\sinh^2 \chi + \sin^2 m}}{\cosh \chi}. \quad (5.48)$$

This is a complex unit, so it acts on wave functions as a unitary operator. We can now write the general solution to the second order difference equation as a sum of solutions to the following first order difference equations,

$$\Psi(t, n; \chi) = \sum_{\epsilon} \Psi_{\epsilon}(t, n; \chi), \quad \Psi_{\epsilon}(t + 1, n; \chi) = U^{\epsilon} \Psi_{\epsilon}(t, n; \chi). \quad (5.49)$$

This is very similar to the time evolution equation in ordinary quantum mechanics,

$$-i\hbar \frac{\partial}{\partial t} \Psi(t) = H \Psi(t) \quad \Leftrightarrow \quad \Psi(t + \tau) = e^{iH\tau/\hbar} \Psi(t). \quad (5.50)$$

The only difference between our time evolution equation and the Schrödinger equation is that we cannot write it as a differential equation, because we do not have a continuous time. We can however use it to read off what the energy of our particle is. For this purpose, we have to write the unitary operator U^{ϵ} , which moves us forward one unit in time, as the exponential of an Hermitian operator,

$$U^{\epsilon} = e^{iH} = \cos H + i \sin H. \quad (5.51)$$

If we compare this to (5.48), we find that H is determined by

$$\cos H = \frac{\cos m}{\cosh \chi}, \quad \epsilon \sin H > 0. \quad (5.52)$$

These are exactly the equations that determine the value of the Euler angle ρ on the mass shell. So we have, on the mass shell, $H = \rho$, and we recover, at the quantum level, exactly what we already found at the classical level in the end of section 2. The energy of the particle is given by the Euler angle ρ , which is related to the deficit angle of the particle, measured in the rest frame of the observer. A peculiar feature of this quantum energy is however that it is only determined up to a multiple of 2π (times the Planck energy), because ρ is an angular variable. Clearly, this is because, at the quantum level, the energy of the particle is actually its *frequency*, and if time is discrete, then this can only be determined up to a multiple of 2π , divided by the time step.

Observables

Finally, we would like to discuss the question how to extract physical information about the particle from the state $|\Psi\rangle$. For this purpose, we need a complete set of *observables*, that is, operators that commute with the mass shell constraint, and for which we have a physical interpretation of the corresponding classical phase space function. This is quite obvious for the operators p_a and J_a , representing the momentum and angular momentum of the particle. They are in fact complete, but not independent. The momentum obeys the mass shell condition, $p_a p^a + \sin^2 m = 0$, and for the angular momentum we have $J_a p^a = 0$. As a set of independent observables we could choose the spatial components p_i and J_i ($i = 1, 2$), but there is another set which has a more natural physical interpretation.

The position operators x^a are not observables as they do not commute with the constraint. But we can use them to construct an observable which represents the position of the particle at a fixed time τ . Classically, this is given by $d_i(\tau) = x_i - p_i/p_0(x_0 + \tau)$, which is easily found to have a vanishing bracket with the constraint. To make it a Hermitian operator, we choose the symmetrized ordering

$$\begin{aligned}\hat{d}_i(\tau) &= \hat{x}_i - (\hat{x}_0 + \tau) \frac{\hat{p}_i}{2\hat{p}_0} - \frac{\hat{p}_i}{2\hat{p}_0} (\hat{x}_0 + \tau) \\ &= \hat{x}_i - \frac{\hat{p}_i}{\hat{p}_0} (\hat{x}_0 + \tau) + i\hbar \frac{\hat{p}_i \hat{u} + \varepsilon_{ij} \hat{p}_j \hat{p}_0}{2\hat{p}_0^2}\end{aligned}\quad (5.53)$$

This looks again similar to the analogue of the relativistic point particle, except the usual “Newton Wigner term” $\frac{1}{2}i\hbar \hat{p}_i/\hat{p}_0^2$ is replaced by a more complicated expression. However, in the low momentum limit, it reduces to the usual one, because we can neglect p_0 compared to $u \approx 1$.

Now, the operators $d_i(\tau)$, for fixed τ , and the spatial components of the momentum p_i form a complete set of independent observables. Classically, they fix the initial condition for the particle at time τ . To set up a position representation, we should diagonalize the observables $d_i(\tau)$. Unfortunately, this is still not possible, because they do not commute. If we consider the operators $d_i(\tau)$ as representing the position of the particle at time τ , then the space on which the particle is moving is non-commutative, like to the non-commutative spacetime that we encountered before we imposed the mass shell condition. To see the similarity, let us compute the commutator of the d_i operators and consider the resulting uncertainty relation. For the classical phase space functions $d_i(\tau)$, we find

$$\{d_i(\tau), d_j(\tau)\} = 2\varepsilon_{ij} \left(\tau + \frac{p_k d_k(\tau)}{p_0} \right). \quad (5.54)$$

Hence, assuming that the same holds up to higher order corrections for the quantum operators, we find that the minimal uncertainty in the position coordinates of a particle centered at the origin grows with the square root of τ . This is exactly the same behaviour as that of the innermost rings in figure 3. So, we can reproduce the same result at the level of observables, that is, with the gauge degrees of freedom completely divided out.

What we cannot reproduce at this level is the discretization of time. This is because there is no observable that corresponds to the measurement of time. Note, for example, that the observable $d_i(\tau)$ is constructed by a kind of interpolation, and it is defined for every real value of τ . The integers do not play any special role. To reproduce the discreteness of time, one should introduce an additional degree of freedom, for example a clock associated with the particle, and then consider the eigenvalues of the operator corresponding to the measurement “read off the clock at external time τ ”. Such an operator is expected to have a discrete spectrum, because it is “canonically conjugate” to the mass of the particle, which is restricted to a finite interval.

Acknowledgements

We would like to thank

Ingemar Bengtsson, Gerard 't Hooft, Jorma Louko and Erik Verlinde

for helpful questions and interesting discussions, and also the referees for helping us to improve this article.

Appendix A The Lorentz group

In this appendix, some basic properties of the Lorentz algebra and its associated Lie groups are collected. The notations is as follows. A vector in three dimensional Minkowski space \mathbb{M}^3 takes a small latin index a, b, c, \dots , running from 0 to 2. The metric is $\eta_{ab} = \text{diag}(-1, 1, 1)$, and the Levi Civita symbol ε^{abc} obeys $\varepsilon^{012} = -\varepsilon_{012} = 1$.

The vector and spinor representation

The group $\text{SO}_+(1, 2)$ of time oriented proper Lorentz transformations acts on a vector in \mathbb{M}^3 from the right as

$$v^a \mapsto v^b u_b^a, \quad \text{where} \quad u_a^b u_c^d \eta_{bd} = \eta_{ac}, \quad u_0^0 > 0. \quad (\text{A.1})$$

Elements of the associated Lie algebra $\mathfrak{so}(1, 2)$ are represented by matrices w_a^b with $w_{ab} = -w_{ba}$. As a vector space, the algebra is isomorphic to \mathbb{M}^3 itself, and its spinor representation is $\mathfrak{sl}(2)$. We denote elements of $\mathfrak{sl}(2)$, that is, traceless 2×2 -matrices, by bold letters. Due to the isomorphism between the $\mathfrak{sl}(2)$ and \mathbb{M}^3 , we can also interpret them as vectors in Minkowski space. Explicitly, the isomorphism is given by expansion in terms of gamma matrices,

$$\boldsymbol{v} = v^a \boldsymbol{\gamma}_a \quad \Leftrightarrow \quad v^a = \frac{1}{2} \text{Tr}(\boldsymbol{v} \boldsymbol{\gamma}^a), \quad (\text{A.2})$$

where

$$\boldsymbol{\gamma}_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \boldsymbol{\gamma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\gamma}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A.3})$$

The algebra of these matrices reads

$$\boldsymbol{\gamma}_a \boldsymbol{\gamma}_b = \eta_{ab} \mathbf{1} - \varepsilon_{abc} \boldsymbol{\gamma}^c, \quad (\text{A.4})$$

where $\mathbf{1}$ is the 2×2 unit matrix. The scalar product on $\mathbb{M}^3 \simeq \mathfrak{sl}(2)$ can be written as $v_a w^a = \frac{1}{2} \text{Tr}(\boldsymbol{v} \boldsymbol{w})$. Timelike, lightlike and spacelike *matrices* \boldsymbol{v} are defined according to the corresponding properties of the *vector* v^a . For example, \boldsymbol{v} is called positive timelike if $\text{Tr}(\boldsymbol{v} \boldsymbol{v}) < 0$ and $\text{Tr}(\boldsymbol{v} \boldsymbol{\gamma}^0) > 0$.

Elements of the group $\text{SL}(2)$ are also denoted by bold letters. An element $\boldsymbol{u} \in \text{SL}(2)$ acts on a vector $\boldsymbol{v} \in \mathfrak{sl}(2)$ in the adjoint representation as

$$\boldsymbol{v} \mapsto \boldsymbol{u}^{-1} \boldsymbol{v} \boldsymbol{u}. \quad (\text{A.5})$$

For this to be the same as (A.1), we must have

$$\mathbf{u}^{-1} \gamma_a \mathbf{u} = u_a^b \gamma_b \quad \Leftrightarrow \quad u_a^b = \frac{1}{2} \text{Tr}(\mathbf{u}^{-1} \gamma_a \mathbf{u} \gamma^b). \quad (\text{A.6})$$

This is the two-to-one group homomorphism mapping the spinor representation $\text{SL}(2)$ onto the vector representation $\text{SO}_+(1, 2)$. The corresponding isomorphism of the algebras $\mathfrak{sl}(2)$ and $\mathfrak{so}(1, 2)$ is given by

$$\mathbf{v} = -\frac{1}{4} \varepsilon^{abc} v_{ab} \gamma_c \quad \Leftrightarrow \quad v_{ab} = \varepsilon_{abc} \text{Tr}(\mathbf{v} \gamma^c). \quad (\text{A.7})$$

As a general convention, we always denote different representations of the same object by the same letter, but with different indices attached, or by a bold letter if it is a 2×2 -matrix.

The group manifold

As a manifold, the group $\text{SL}(2)$ is naturally embedded in \mathbb{R}^4 (or $\mathbb{R} \times \mathbb{M}^3$). The most convenient embedding coordinates are obtained by expanding \mathbf{u} in terms of the unit and the gamma matrices. We define a scalar u and a vector p_a such that

$$\mathbf{u} = u \mathbf{1} + p_a \gamma^a \quad \Rightarrow \quad u^2 - p_a p^a = 1, \quad (\text{A.8})$$

where the last equations results from the condition that $\det \mathbf{u} = 1$. It defines a $(2, 2)$ -hyperboloid embedded in \mathbb{R}^4 , which is shown in figure 1. We can also define a “projection” of the group into the algebra,

$$\mathbf{u} = u \mathbf{1} + p_a \gamma^a \quad \mapsto \quad \mathbf{p} = p_a \gamma^a. \quad (\text{A.9})$$

The image of the group manifold under this projection is the set of all vectors $\mathbf{p} \in \mathfrak{sl}(2)$ with length squared bigger or equal to -1 . It is the subset of Minkowski space between the two unit timelike hyperboloids.

A useful set of coordinates on the group manifold is provided by the “Euler angles” (ρ, χ, ϕ) . The general solution to (A.8) can be written as

$$\begin{aligned} u &= \cos \rho \cosh \chi, & p_1 &= \cos \phi \sinh \chi, \\ p^0 &= \sin \rho \cosh \chi, & p_2 &= \sin \phi \sinh \chi. \end{aligned} \quad (\text{A.10})$$

The two “angular” coordinates ρ and ϕ have a period of 2π , and χ is a “hyperbolic angle” with $\chi \geq 0$. For $\chi = 0$, the angle ϕ becomes redundant, whereas ρ is always regular. All together, they form a kind of “cylindrical” coordinate system, with radial coordinate χ , polar angle ϕ , and ρ playing the role of the “ z ”-coordinate, which is however winded up. The coordinates ρ and χ are shown as the grid lines in figure 1. They are called Euler angles because they are closely related to the Euler angles of $\text{SU}(2)$. We can write the group element corresponding to (A.10) as

$$\mathbf{u} = e^{\frac{1}{2}(\rho+\phi)\gamma_0} e^{\chi\gamma_1} e^{\frac{1}{2}(\rho-\phi)\gamma_0}. \quad (\text{A.11})$$

The Lorentz transformation represented by \mathbf{u} is decomposed into a rotation of the spatial plane by $\rho + \phi$, a boost in the γ_2 direction with parameter χ , and another rotation

by $\rho + \phi$. To see this, and to check that (A.11) is actually the same as (A.10), one has to use of the following formulas for the exponential map $\mathfrak{sl}(2) \rightarrow \text{SL}(2)$,

$$\begin{aligned} e^{\alpha\gamma_0} &= \cos \alpha \mathbf{1} + \sin \alpha \gamma_0, \\ e^{\beta\gamma_1} &= \cosh \beta \mathbf{1} + \sinh \beta \gamma_1. \end{aligned} \quad (\text{A.12})$$

They imply that the exponential of γ_0 generates a rotation in the γ_1 - γ_2 plane,

$$\begin{aligned} e^{-\frac{1}{2}\alpha\gamma_0} \gamma_1 e^{\frac{1}{2}\alpha\gamma_0} &= \cos \alpha \gamma_1 + \sin \alpha \gamma_2, \\ e^{-\frac{1}{2}\alpha\gamma_0} \gamma_2 e^{\frac{1}{2}\alpha\gamma_0} &= \cos \alpha \gamma_2 - \sin \alpha \gamma_1, \end{aligned} \quad (\text{A.13})$$

and similarly, that the exponential of γ_1 generates a boost in γ_2 direction and vice versa.

Rotations and the mass shells

The projection (A.9) can be used to distinguish between timelike, lightlike, and spacelike *group elements* \mathbf{u} , according to the properties of the vector \mathbf{p} . The timelike ones are then those with $-1 < u < 1$. On the boundaries of this interval we have the lightlike group elements, and outside this range the spacelike ones. The vector \mathbf{p} points along the “axis” of the Lorentz transformation represented by \mathbf{u} , which can be seen from the fact that it is invariant under the adjoint action,

$$\mathbf{u}^{-1} \mathbf{p} \mathbf{u} = \mathbf{p}. \quad (\text{A.14})$$

If this axis is timelike, then the Lorentz transformation is a *rotation* of the spacelike plane orthogonal to \mathbf{p} . If it is spacelike, then it is a *boost*, and in between we have the lightlike Lorentz transformations. The rotations play a special role, because they appear as the momenta of massive particles. The mass of the particle is related to the angle of rotation by $\alpha = 2m$. To find all group elements representing a rotation by $2m$, consider a unit spacelike vector \mathbf{v} orthogonal to the axis \mathbf{p} , so that

$$\mathbf{v}\mathbf{v} = \mathbf{1}, \quad \mathbf{v}\mathbf{p} + \mathbf{p}\mathbf{v} = 0 \quad \Rightarrow \quad \mathbf{u}^{-1}\mathbf{v} = \mathbf{v}\mathbf{u}. \quad (\text{A.15})$$

Under the rotation, \mathbf{v} is mapped onto $\mathbf{u}^{-1}\mathbf{v}\mathbf{u}$. If this is a rotation by $2m$, then the scalar product of this vector with \mathbf{v} must be $\cos(2m)$, independent of the choice of \mathbf{v} . Hence, we must have

$$\frac{1}{2}\text{Tr}(\mathbf{v}\mathbf{u}^{-1}\mathbf{v}\mathbf{u}) = \frac{1}{2}\text{Tr}(\mathbf{v}\mathbf{v}\mathbf{u}\mathbf{u}) = \frac{1}{2}\text{Tr}(\mathbf{u}\mathbf{u}) = \cos(2m). \quad (\text{A.16})$$

If we expand \mathbf{u} in terms of the gamma matrices according to (A.8), this becomes

$$u^2 - 1 = \cos(2m) \quad \Leftrightarrow \quad u^2 = \cos^2 m. \quad (\text{A.17})$$

The solutions to this equation fall into four conjugacy classes of $\text{SL}(2)$, containing the elements $\pm e^{\pm m\gamma_0}$. The two sign ambiguities arise as follows. The sign in the exponent fixes the *direction* of the rotation. In contrast to three dimensional Euclidean space, in

Minkowski space we can distinguish between rotations in positive and negative direction. The second sign appears because \mathbf{u} is an element of the *spinor* representation of the Lorentz group, whereas “rotation by $2m$ ” refers to the *vector* representation. Thus, for every rotation we have two spinor representations, which differ by an overall sign.

To get rid of at least one sign, we change the definition of “rotation by an angle $2m$ ” slightly, such that it also fixes the action on spinors. In the vector representation, we cannot distinguish a rotation in the positive direction by $2m > 0$ from a rotation in the negative direction by $2m - 2\varphi < 0$. But in the spinor representation we can. The former is then represented by an element in the conjugacy class of $e^{m\gamma_0}$, whereas the second lies in the class of $e^{(m-\pi)\gamma_0} = -e^{m\gamma_0}$. So, if we only call the former a “rotation by $2m$ ”, then the condition becomes

$$\mathbf{u} = \cos m \quad \Leftrightarrow \quad \frac{1}{2}\text{Tr}(\mathbf{u}) = \cos m. \quad (\text{A.18})$$

As this equation is still invariant under $m \mapsto -m$, or $\mathbf{u} \mapsto \mathbf{u}^{-1}$, it still does not fix the direction of the rotation. This ambiguity remains in the description of the particle, and it becomes a kind of particle–antiparticle duality.

Appendix B Harmonic analysis

Here, we shall briefly derive the basic results of the harmonic analysis on the group manifold $\text{SL}(2)$. For a more detailed treatment of the subject we refer to [20, 21, 22, 23]. We start with the general $T = x^0$ and $M = J^0$ eigenstates $|t, n; \psi\rangle$, where ψ is function the Euler angle χ , which is essentially the spatial momentum and commutes with T and M (see (5.17)). The problem is to diagonalize the operator $S^2 = x^a x_a$ on these states, which is equal to the Laplace Beltrami operator. Explicitly,

$$\widehat{S}^2 |t, n; \psi\rangle = |t, n; \Delta\psi\rangle, \quad (\text{B.1})$$

where

$$\Delta = \hbar^2 \left(\frac{1}{\sinh(2\chi)} \frac{\partial}{\partial \chi} \sinh(2\chi) \frac{\partial}{\partial \chi} - \frac{4n}{\sinh^2(2\chi)} + \frac{2nt + t^2}{\cosh^2 \chi} \right). \quad (\text{B.2})$$

It is Hermitian with respect to the scalar product

$$\langle t, n; \psi | t', n'; \phi \rangle = \delta_{t,t'} \delta_{n,n'} \int_0^\infty d\chi \, 2 \sinh(2\chi) \, \psi^*(\chi) \phi(\chi). \quad (\text{B.3})$$

Let us introduce the following raising and lowering operators, which change the quantum numbers t and n ,

$$T_\pm = x_1 \pm ix_2, \quad M_\pm = J_1 \mp iJ_2 + \frac{1}{2}(x_1 \mp ix_2). \quad (\text{B.4})$$

They act on \mathbf{u} eigenstates as

$$\widehat{T}_\pm |\mathbf{u}\rangle = 2i\hbar \mathcal{L}_\pm |\mathbf{u}\rangle, \quad \widehat{M}_\pm |\mathbf{u}\rangle = i\hbar \mathcal{R}_\mp |\mathbf{u}\rangle, \quad (\text{B.5})$$

where $\mathcal{L}_\pm = \mathcal{L}_1 \pm i\mathcal{L}_2$ and $\mathcal{R}_\pm = \mathcal{R}_1 \pm i\mathcal{R}_2$ are combinations of the left and right invariant vector fields. Using this and $\mathcal{L}_-\mathcal{L}_+ = \mathcal{L}_a\mathcal{L}^a + \mathcal{L}^0\mathcal{L}^0 + i\mathcal{L}^0$, one finds the following relations:

$$\begin{aligned}\widehat{T}_+\widehat{T}_- &= \widehat{S}^2 + \widehat{T}(\widehat{T} - 2\hbar), \\ \widehat{T}_-\widehat{T}_+ &= \widehat{S}^2 + \widehat{T}(\widehat{T} + 2\hbar).\end{aligned}\tag{B.6}$$

Similarly, with the same expressions for the right invariant vector fields, one obtains

$$\begin{aligned}\widehat{M}_+\widehat{M}_- &= \frac{1}{4}\widehat{S}^2 + (\widehat{M} + \frac{1}{2}\widehat{T})(\widehat{M} + \frac{1}{2}\widehat{T} - \hbar), \\ \widehat{M}_-\widehat{M}_+ &= \frac{1}{4}\widehat{S}^2 + (\widehat{M} + \frac{1}{2}\widehat{T})(\widehat{M} + \frac{1}{2}\widehat{T} + \hbar).\end{aligned}\tag{B.7}$$

We also need the commutator relations

$$\begin{aligned}[\widehat{T}, \widehat{T}_\pm] &= \pm 2\hbar \widehat{T}_\pm, & [\widehat{T}, \widehat{M}_\pm] &= 0, \\ [\widehat{M}, \widehat{T}_\pm] &= \mp \hbar \widehat{T}_\pm, & [\widehat{M}, \widehat{M}_\pm] &= \pm \hbar \widehat{M}_\pm.\end{aligned}\tag{B.8}$$

Now, let $|t, n, s^2\rangle$ be a normalized eigenstate of T , M , and S^2 , with eigenvalues $\hbar t$, $\hbar n$, and $\hbar^2 s^2$, respectively. Acting on this with the raising and lowering operators above, we can increase and decrease the quantum numbers t and n ,

$$\widehat{T}_\pm |t, n, s^2\rangle \propto |t \pm 2, n \mp 1, s^2\rangle, \quad \widehat{M}_\pm |t, n, s^2\rangle \propto |t, n \pm 1, s^2\rangle.\tag{B.9}$$

If we compute the norm of these states, we find

$$\begin{aligned}\|\widehat{T}_\pm |t, n, s^2\rangle\| &= \langle t, n, s^2 | \widehat{T}_\mp \widehat{T}_\pm |t, n, s^2\rangle \\ &= \langle t, n, s^2 | (\widehat{S}^2 + \widehat{T}(\widehat{T} \pm 2\hbar)) |t, n, s^2\rangle \\ &= \hbar^2 (s^2 + t(t \pm 2)). \\ \|\widehat{M}_\pm |t, n, s^2\rangle\| &= \langle t, n, s^2 | \widehat{M}_\mp \widehat{M}_\pm |t, n, s^2\rangle \\ &= \langle t, n, s^2 | (\frac{1}{4}\widehat{S}^2 + (\widehat{M} + \frac{1}{2}\widehat{T})(\widehat{M} + \frac{1}{2}\widehat{T} \pm \hbar)) |t, n, s^2\rangle \\ &= \frac{1}{4}\hbar^2 (s^2 + (2n + t)(2n + t \pm 2)).\end{aligned}\tag{B.10}$$

These have to be non-negative, which gives the following algebraic restrictions on the quantum numbers,

$$\begin{aligned}s^2 + t(t + 2) &\geq 0 \quad (\widehat{T}_+), & s^2 + (2n + t)(2n + t + 2) &\geq 0 \quad (\widehat{M}_+), \\ s^2 + t(t - 2) &\geq 0 \quad (\widehat{T}_-), & s^2 + (2n + t)(2n + t - 2) &\geq 0 \quad (\widehat{M}_-).\end{aligned}\tag{B.11}$$

Moreover, we infer that starting from any state $|t, n, s^2\rangle$, we can build up a two dimensional ‘‘tower’’ of states by acting with the raising and lowering operators on them. We can reach any value for n and t , except that we cannot change the parity of t , and provided that the series does not terminate. It will only terminate if one of these inequalities is saturated. The operators which annihilate these states are then given in parenthesis behind the inequalities. Whether the inequalities are satisfied or not depends crucially on the range in which s^2 lies. Let us distinguish the following cases:

- for $s^2 \geq 1$, the spectrum will be continuous,
- for $0 < s^2 < 1$, there will be no states at all,
- and for $s^2 \leq 0$, there will be a discrete series of eigenstates.

Let us start with the case $s^2 \geq 1$, and introduce a new quantum number $\lambda \in \mathbb{R}_+$ such that $s^2 = \lambda^2 + 1$. The left hand sides of (B.11) are now always positive, and we do not get any algebraic restrictions on the quantum numbers. But can we realize these states as square integrable functions on $\text{SL}(2)$? If we write the eigenstate as $|t, n, s^2\rangle = |t, n; \psi\rangle$, with a wave function $\psi(\chi)$, then we can study the behaviour of ψ for $\chi \rightarrow \infty$. For large χ , the eigenvalue equation of the Laplace Beltrami operator (B.2) becomes

$$\begin{aligned} e^{-2\chi} \frac{\partial}{\partial \chi} e^{2\chi} \frac{\partial}{\partial \chi} \psi(\chi) &\sim (\lambda^2 + 1) \psi(\chi) \\ \Rightarrow \psi(\chi) &\sim A e^{-(1-i\lambda)\chi} + B e^{-(1+i\lambda)\chi}. \end{aligned} \quad (\text{B.12})$$

We get a superposition of an ingoing and outgoing wave with wavelength $1/\lambda$. They fall off with $e^{-\chi}$. This is exactly the right behaviour for ψ to become normalizable under the scalar product (B.3), like usual plane waves in $L_2(\mathbb{R})$. As χ is a radial coordinate, there must be a balance between ingoing and outgoing waves, and instead of two we only get one solution for ψ for any fixed set of quantum numbers t, n and s^2 . So we have the following continuous series of eigenstates,

$$\hat{S}^2 |t, n, \lambda\rangle = \hbar^2 (\lambda^2 + 1) |t, n, \lambda\rangle, \quad t, n \in \mathbb{Z}, \quad \lambda \in \mathbb{R}_+. \quad (\text{B.13})$$

Next we consider the case $0 < s^2 < 1$. This time we repack s^2 by a real number l , such that $s^2 = l(2 - l)$, and for the moment we take $1 < l < 2$. We then get some algebraic restrictions on the quantum numbers from (B.11). They are violated if $t = \pm 1$ or $t + 2n = \pm 1$. So we cannot have any states with $t = \pm 1$. But the inequalities are never saturated, and so we can change t by any multiple of two, using the T_{\pm} operators. This excludes all odd values for t . On the other hand, if there is an eigenstate $|t, n, s^2\rangle$ for some even t , then we can use the raising and lowering operations to arrive at a state $|0, 0, s^2\rangle = |0, 0; \psi\rangle$, and the wave function ψ has to obey

$$\psi''(\chi) + 2 \coth(2\chi) \psi'(\chi) = l(2 - l) \psi(\chi). \quad (\text{B.14})$$

It is then not difficult to show that a solution to this differential equation can never be square integrable and smooth at the origin $\chi = 0$. So there are no states with eigenvalues $0 < s^2 < 1$.

Finally, let us study the case $s^2 \leq 0$. Here we also use the parameter l such that $s^2 = l(2 - l)$, but now we take $l \geq 2$. Then we find the following condition on the quantum numbers,

$$t(t + 2) \geq l(l - 2) \quad \text{and} \quad t(t - 2) \geq l(l - 2). \quad (\text{B.15})$$

As $l \geq 2$, this is equivalent to the somewhat simpler condition

$$t \geq l \quad \text{or} \quad t \leq -l, \quad (\text{B.16})$$

Now assume that l is *not* an integer. Then these inequalities will never be saturated, which in turn means that starting from any state we can build up an infinite tower of states in both t -directions, by acting on it with the T_{\pm} operators. Thereby, we can increase or decrease t by any multiple of 2. However, at least one of these values will lie in the region between l and $-l$, which is forbidden. So l has to be an integer. But even then we can create towers of states with elements falling into the forbidden region, unless the series terminates.

The only way to build up such a terminating tower is to start on the boundary, with quantum number $t = l$ or $t = -l$, and then act on this state with T_+ or T_- , respectively. Note that it is T_- that annihilates the state at $t = l$ and vice versa, so what we get are two infinite towers, one with $t = l, l + 2, l + 4, \dots$ and the other one with $t = -l, -l - 2, -l - 4, \dots$. For a fixed t , we therefore get the restriction that

$$2 \leq l \leq |t| \quad \text{and} \quad t \equiv l \pmod{2}. \quad (\text{B.17})$$

There are no states for $t = 0$ and $t = \pm 1$. With a similar argument, we can determine the range of n for fixed l and t . The result is

$$n \geq \frac{1}{2}(l - t) \quad \text{for} \quad t \geq 2, \quad n \leq -\frac{1}{2}(l - t) \quad \text{for} \quad t \leq -2. \quad (\text{B.18})$$

Hence, for each value of $l = 2, 3, 4, \dots$, we have two “ground states” $|l, 0, l\rangle$ and $|-l, 0, l\rangle$, and all other states can be generated by acting with T_+ and M_+ onto $|l, 0, l\rangle$, and similarly with the T_- and M_- onto $|-l, 0, l\rangle$. Note that thereby also negative values for n occur in the positive series and vice versa, because T_+ *decreases* n and T_- *increases* it. Studying the asymptotic behaviour and of these states, we find that

$$\psi(\chi) \sim A e^{-l\chi} + B e^{l\chi}. \quad (\text{B.19})$$

For this to be square integrable, we must have $B = 0$. Thus, there is again only one state for each set of quantum numbers, and the whole discrete series of eigenstates is given by

$$\hat{S}^2 |t, n, l\rangle = \hbar^2 l(2 - l) |t, n, l\rangle, \quad l = 2, 3, 4, \dots \quad (\text{B.20})$$

which splits into a positive and a negative series with

$$\begin{aligned} t &= l, l + 2, l + 4, \dots, & n &\geq \frac{1}{2}(l - t), \\ t &= -l, -l - 2, -l - 4, \dots, & n &\leq -\frac{1}{2}(l - t). \end{aligned} \quad (\text{B.21})$$

References

- [1] A. Staruszkiewicz: Gravitation theory in three-dimensional space, *Acta Phys. Polon.* **24** (1963) 734.
- [2] S. Deser, R. Jackiw and G. 't Hooft: Three-dimensional Einstein gravity: Dynamics of flat space, *Ann. Phys.* **152** (1984) 220.

- [3] A. Achúcarro and P. Townsend: A Chern-Simons action for three-dimensional anti-de-Sitter supergravity theories, *Phys. Lett.* **B180** (1986) 85.
- [4] E. Witten: 2+1 Dimensional gravity as an exactly soluble system, *Nucl. Phys.* **B331** (1988) 46.
- [5] S. Carlip: Six ways to quantize (2+1)-dimensional gravity, gr-qc/9305020.
- [6] S. Carlip: Exact quantum scattering in 2+1 dimensional gravity, *Nucl. Phys.* **B324** (1989) 106.
- [7] G. 't Hooft: Causality in (2+1)-dimensional gravity, *Class. Quantum Grav.* **9** (1992) 1335.
- [8] G. 't Hooft: Nonperturbative two particle scattering amplitudes in (2+1)-dimensional quantum gravity, *Comm. Math. Phys.* **117** (1988) 685.
- [9] G. 't Hooft: Quantization of point particles in (2+1)-dimensional gravity and spacetime discreteness, *Class. Quantum Grav.* **13** (1996) 1023.
- [10] G. 't Hooft: Canonical quantization of gravitating point particles in 2+1 dimensions, *Class. Quantum Grav.* **10** (1993) 1653.
- [11] A. Bellini, M. Ciafaloni and P. Valtancoli: Solving the N-body problem in (2+1)-gravity, *Nucl. Phys.* **B462** (1996) 453.
- [12] H. Snyder: Quantized space-time, *Phys. Rev.* **71** (1947) 38.
- [13] H.J. Matschull: Phase space structure of multi-particles models in 2+1 gravity, in preparation.
- [14] H.J. Matschull: Three-dimensional canonical quantum gravity, *Class. Quantum Grav.* **12** (1995) 2621.
- [15] M. Welling: Two particle quantum mechanics in 2+1 Gravity using non commuting coordinates, gr-qc/9703058.
- [16] M. Welling: Gravity in 2+1 dimensions as a Riemann-Hilbert problem, *Class. Quantum Grav.* **13** (1996) 653.
- [17] A. Demichev: Physical meaning of quantum space time symmetries, hep-th/9701079.
- [18] G. Grignani and G. Nardelli: Gravity and the Poincaré group, *Phys. Rev.* **D45** (1992), 2719.
- [19] P. Salomonson, B.S. Skagerstam and A. Stern: ISO(2,1) chiral models and quantum gravity in (2+1)-dimensions, *Nucl. Phys.* **B347** (1990), 769.
- [20] N. Limić, J. Niederle and R. Raczka: Discrete (Continuous) degenerate representations of non compact rotation groups I & II, *J. Math. Phys.* **7** (1966) 1861 & 2026.

- [21] R. Dijkgraaf, H. Verlinde and E. Verlinde: String propagation in a black hole geometry, *Nucl. Phys.* **B371** (1992) 269.
- [22] W.J. Holman and L.C. Biedenharn: Complex angular momenta and the groups $SU(1,1)$ and $SU(2)$, *Ann. Phys.* **39** (1966) 1.
- [23] N.J. Vilenkin and A.U. Klimyk: *Representation of Lie Groups and Special Functions Vol. 1*, (Kluwer, Dordrecht, 1991).